IES Institute of Education Sciences

Evaluation Report

Study of Teacher Coaching Based on Classroom Videos: Impacts on Student Achievement and Teachers' Practices

NCEE 2022-006r
U.S. Department of Education

A Publication of the National Center for Education Evaluation at IES

The Institute of Education Sciences (IES) is the independent, non-partisan statistics, research, and evaluation arm of the U.S. Department of Education. The IES mission is to provide scientific evidence on which to ground education practice and policy and to share this information in formats that are useful and accessible to educators, parents, policymakers, researchers, and the public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other IES product or report, we would like to hear from you. Please direct your comments to ncee.feedback@ed.gov.

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-16-C-0021 by Mathematica. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

June 2022

Clark, M., Max, J., James-Burdumy, S., Robles, S., McCullough, M., Burkander, P., and Malick, S. (2022). Study of Teacher Coaching Based on Classroom Videos: Impacts on Student Achievement and Teachers' Practices (NCEE 2022-006r). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

This report is available on the Institute of Education Sciences website at http://ies.ed.gov/ncee.

IES Institute of Education Sciences

Study of Teacher Coaching Based on Classroom Videos: Impacts on Student Achievement and Teachers' Practices

June 2022

Melissa Clark
Jeffrey Max
Susanne James-Burdumy
Silvia Robles
Moira McCullough
Paul Burkander
Steven Malick
Mathematica

NCEE 2022-006r
U.S. DEPARTMENT OF EDUCATION

Helping teachers become more effective in the classroom is a high priority for educators and policymakers. A growing body of evidence suggests that individualized coaching focused on general teaching practices can improve teachers' instruction and student achievement. However, little is known about the benefits of specific approaches to coaching, including who is doing the coaching, how coaches observe teachers' instruction, and how or how often coaches provide feedback to teachers. This study examined one promising strategy for individualized coaching: professional coaches—rather than district or school staff—providing feedback to teachers based on videos of their instruction. Feedback based on videos gives teachers the opportunity to observe and reflect on their own teaching and allows coaches to show teachers specific moments from their teaching when providing feedback. For this study, about 100 elementary schools were randomly divided into three groups: one that received fewer highly structured cycles of focused professional coaching during a single school year (five cycles), one that received more (eight cycles), and one that continued with its usual strategies for supporting teachers. The study compared teachers' experiences and student achievement across the three groups to determine the effectiveness of the two versions of the coaching.

Key Findings

- *Five coaching cycles based on videos of teachers' instruction improved students' achievement, including for novice teachers and those with weaker classroom practices at the start of the study.*
- *Eight cycles of coaching was not effective. Eight cycles of the coaching did not affect student achievement, perhaps because teachers had less time during each cycle to work on the practices being addressed.*
- *The study's coaching changed the type of feedback that teachers received. Compared to those who did not receive the study's coaching, teachers who received the coaching were more likely to report receiving feedback that focused on specific teaching practices, included strategies to use in their classrooms, and provided opportunities to observe and reflect on their teaching.*

Teachers play a critical role in students' learning and long-term success in school and the labor market.[1] As a result, states, districts, and the federal government make substantial investments in time and resources to develop teachers' skills.[2] However, educators and policymakers lack strong evidence about the specific features of professional development that are critical for improving student achievement.[3] This evidence is particularly important as states and districts seek strategies to address the COVID-19 pandemic's detrimental effects on student learning.

To build evidence on how best to support teachers, this study examined the effectiveness of one promising approach—ongoing, remote feedback from professional coaches focused on teachers' general classroom practices, based on videos of the teachers' instruction. Several studies suggest that coaching can improve student achievement, including a few small-scale studies focused on the approach to coaching examined in this study.[4] But the evidence does not provide clear guidance on how coaches should provide feedback to teachers, how much coaching teachers should receive, or whether watching videos of their own teaching can help teachers improve.[5] Teachstone, a professional development provider, was selected through a national competition to provide the study's coaching. Its selection was based in part on prior evidence of the program's effectiveness for improving teachers' practices and student learning.[6] Replicating the program on a larger scale and varying the number of coaching cycles would add to the information about the effectiveness of its approach.

Providing feedback based on video recordings of teachers' lessons offers a few potential benefits. Coaches can use videos of teachers' classrooms to show teachers specific aspects of their teaching that are working well or that need to be improved. Videos also allow teachers to reflect on their practices in new ways—for example, to see aspects of their teaching or students' reactions to their teaching they had not noticed before or to think about why a teaching practice did or did not work.[7] In addition, using video-based observations allows coaches to provide feedback remotely, providing more flexibility for districts to hire professional coaches who are located outside the district's immediate geographic area.

Providing feedback on general, rather than subject-specific, teaching practices also offers potential benefits. In particular, it may be more feasible and cost-effective for districts. Rather than hiring separate coaches for different subject areas, districts or schools can hire a single coach to provide teachers feedback across subject areas.[8]

To shed light on how much coaching is needed to improve teacher effectiveness, the study tested two versions of the coaching that both occurred over a single school year—a version that provided five cycles of coaching and a version that provided eight cycles. Coaches were expected to complete a cycle every three weeks for teachers in the five-cycle group and every two weeks for teachers in the eight-cycle group (to ensure coaches had sufficient time to complete all eight cycles by the end of the school year). A comparison of the two versions thus reflects the effects of five less-condensed cycles of coaching versus eight more-condensed cycles—a tradeoff that schools might consider when deciding how much coaching to provide in a school year. The same coaches delivered both versions, but with the five-cycle version in one set of schools and the eight-cycle version in another set of schools.

The study's coaching had two additional key features:

Feedback on a targeted set of practices that is tailored to teacher needs. Study coaches focused their feedback on a set of teaching practices from a well-established classroom observation tool.[9] The practices were grouped into three broad areas: classroom management, building supportive relationships with students, and building students' understanding of the content being taught. Over time, coaches were expected to cover practices across all three of these areas, tailoring the coaching to each teacher's needs. To do so, they worked with teachers to select two to three specific practices to address in each cycle, based on both teachers' interests and areas in which the coach thought the teacher could improve. As teachers became more adept at practices related to classroom management and building supportive relationships with students, the coaching was expected to focus more exclusively on practices related to building students' understanding of content. However, coaches had discretion to deviate from this expected sequence based on teachers' individual needs.

Structured approach to the coaching. The coaching used a well-defined, structured approach to providing feedback to teachers. Coaches delivered the coaching in cycles that consisted of the five steps shown in Exhibit 1. The coaches received training and detailed guidance on how to implement each step of the cycle:

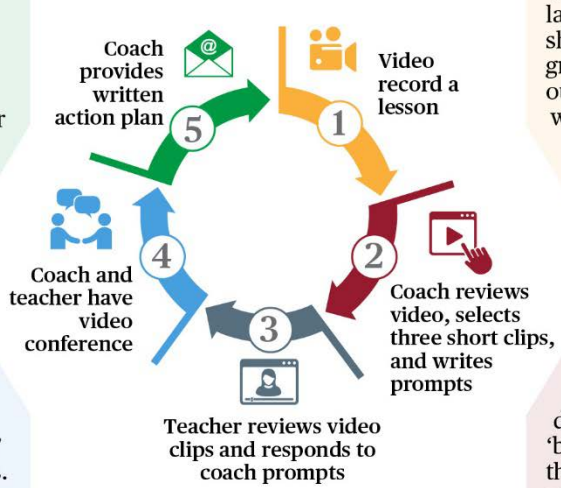
- **Step 1: Video record a lesson.** The teacher and coach identified a lesson to video record that provided opportunities for the teacher to use the targeted practices being addressed in the coaching cycle.
- **Step 2: Coach reviews video, selects three short clips, and writes prompts.** Coaches were trained to select three video clips from a teacher's video-recorded lesson for each coaching session. Each clip focused on one or two of the targeted practices for that cycle and had a specific purpose: the first clip showed positive aspects of how the teacher used the targeted practice; the second clip showed how the teacher's use of the targeted practice influenced students' behavior; and the third clip showed how the teacher's use of the targeted practice led to student learning. Coaches also followed a specific structure for writing prompts to

help teachers reflect on their video clips: they would name the practice that was the focus of the clip, describe the practice, and ask a question about the teacher's use of the practice in the clip.

- **Step 3: Teacher reviews video clips and responds to coach prompts.** Teachers were expected to watch their video clips and provide written responses to each question posed by the coach about the clips. Coaches reviewed these responses in preparation for their conference together.
- **Step 4: Coach and teacher have video conference**. Within each cycle the coach held a conference with each teacher that followed a specific structure. Coaches first checked to see how teachers were doing to help build rapport. They then discussed and provided feedback on the targeted teaching practices, video clips, and the teacher's written responses to the coach's prompts. They next discussed detailed strategies for improving these practices. Finally, they worked with the teacher to select teaching practices to address in the next cycle.
- **Step 5: Coach provides written action plan**. After the conference, the coach provided the teacher an action plan that contained a short description of the targeted practices, web links to one or two videos of exemplar teachers using the practices, and a note about the lesson the teacher would video record for the next cycle.

Teachers were expected to spend about 75 minutes over the course of each cycle reviewing clips, responding to the coach's prompts, and meeting with their coach, in addition to any time working independently on the cycle's focal practices.

Exhibit 1. Steps in each coaching cycle



Example action plan: The plan summarizes the conference and activities for the next cycle, including the next video recording: a math lesson where the teacher encourages students to ask each other questions and expand on other students' ideas. The coach assigns a video of an exemplar teacher using the practice and one page from a guidebook describing it.

Example lesson: For a coaching cycle focused on asking students follow-up questions to deepen understanding, the teacher video records an English language arts lesson where she works with a small group of students to figure out the meaning of new words when reading.

Example conference: The teacher and coach discuss how asking open-ended questions and expanding on students' responses improved students' participation in conversations. The coach notes that the teacher could build on this by encouraging students to ask each other questions and make connections to what other students said. They decide to focus on this practice in the next cycle.

Example prompt: "In this clip, you had students find clues in the text to help them determine the meaning of 'biological.' You asked them the meaning of the root words 'bio' and 'logi,' asked what 'biological' might mean, and helped them connect this to context clues in the text. What did students do or say to show your feedback loops were effective?"

Example response: "Students answered my questions about the meaning of 'biological' by remembering context clues or referring to their list of root words. The students better understood the value of context clues once they knew the word's meaning."

Box 1. The study design

Who participated?

- Schools: 107 elementary schools from 14 large urban and suburban districts that were not already providing extensive professional coaching and feedback to teachers
- Teachers: 353 4th- and 5th-grade teachers who taught math, English language arts, or both; with an average of 12 years of teaching experience
- Students: 8,906 students in 4th and 5th grades in study schools in the 2018-2019 school year
  - 67 percent were eligible for the federal free or reduced-price lunch program
  - 66 percent were students of color
  - 10 percent were English learners
  - 10 percent received special education services

How was the study conducted?

- Random assignment: Each school chose whether its 4th- or 5th-grade teachers would participate in the study. Within each district, schools with similar characteristics were then grouped into sets of three schools. Within each set of three schools, one school was randomly assigned to a group whose teachers were to receive five coaching cycles in the 2018-2019 school year, one to a group whose teachers were to receive eight coaching cycles in the 2018-2019 school year, and one to a group whose teachers did not receive any coaching from the study. Each coach was randomly assigned a roughly equal number of teachers from the five- and eight-cycle groups to ensure that the same coaches delivered the coaching to both groups.
- Analysis: The analysis compared outcomes across the three groups of schools after one year of coaching to measure the effects of five and eight cycles of the coaching on teachers' practices and student achievement.

What data were used?

- Student achievement: State assessment test scores in math and English language arts from the 2018-2019 school year as well the 2017-2018 school year to account for student achievement prior to the study coaching in the analysis
- Teachers' practices: Ratings of teachers' classroom instruction during the 2018-2019 school year using the Classroom Assessment Scoring System (CLASS) to measure teachers' general teaching practices
- Implementation of the coaching: Data from Teachstone's online coaching platform documenting activities covered in each coaching cycle

FIVE CYCLES OF COACHING BASED ON CLASSROOM VIDEOS IMPROVED STUDENT ACHIEVEMENT

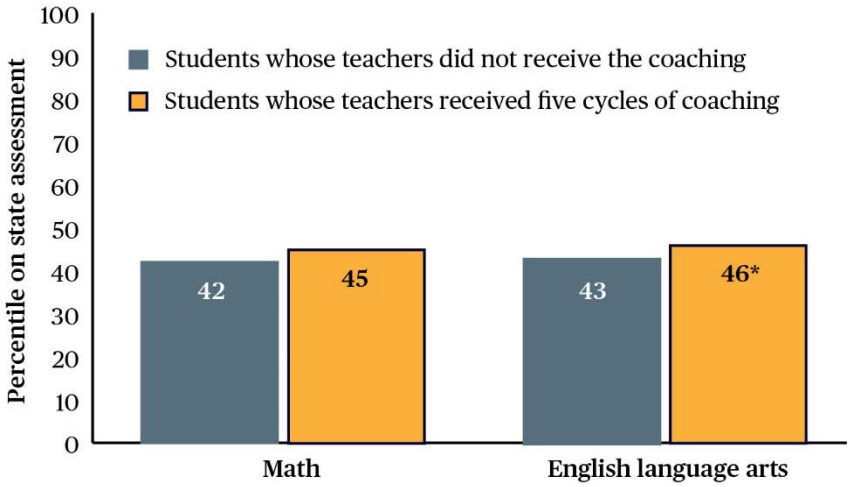
Given the competing demands on teachers' time and district resources, the study examined the effectiveness of five cycles of coaching, which was considered easier for districts to implement than eight cycles and yet still sufficient to generate changes in teacher practice that could influence student achievement. A few other studies suggest that providing individualized feedback to teachers with similarly low frequency during a school year can improve student achievement, although teachers in these studies received feedback that was less structured than the study's coaching.[10] Exhibit 2 provides an example of practices a teacher and coach might have focused on during each cycle, across the three broad areas of teaching practices covered by the coaching.

Exhibit 2. Example of practices a teacher and coach might have focused on during each coaching cycle, across the three broad areas of teaching practices covered by the coaching

	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5
Strategies to build students' understanding of the content	*The coach and teacher use the first cycle to get to know each other and discuss a plan for the coaching*	• Present learning targets for lessons • Pair up students to share thoughts about why learning targets are important	• Help students connect their prior knowledge to the new content they learn • Address students' misconceptions about the content	• Model your thinking for students and help students verbalize their own thought processes as they work • Build on students' responses to questions by asking follow-up questions	• Have students develop arguments based on evidence and explain their answers to peers • Have students write down questions to ask other students during a discussion • Have students build on the responses of other students when they agree or disagree
Strategies to improve classroom management			• Quickly redirect students who are not on task		
Strategies to build supportive relationships with students		• Connect lesson content to students' lives outside school		• Affirm students' efforts during class and provide time and space for students to work through tasks and develop their own solutions	

- ***Five cycles of coaching improved student achievement in English language arts, by an amount equivalent to three percentile points on state assessments or about two additional months of learning on average.*** Students taught by teachers in the five-cycle coaching group had higher English language arts test scores at the end of the school year than students taught by teachers who did not receive the coaching (Exhibit 3).[11] Average English language arts scores for students of teachers in the five-cycle group were at the 46th percentile on the state assessment, meaning they scored higher than 46 percent of all students in the state who took the test. This compared with average scores at the 43rd percentile for students whose teachers did not receive the coaching. This difference is equal to almost two additional months of learning.[12] Although the coaching resulted in a similar difference in students' average math scores, the study could not definitively conclude that it improved math achievement for students overall.[13]

Exhibit 3. Effects of five cycles of coaching on student achievement



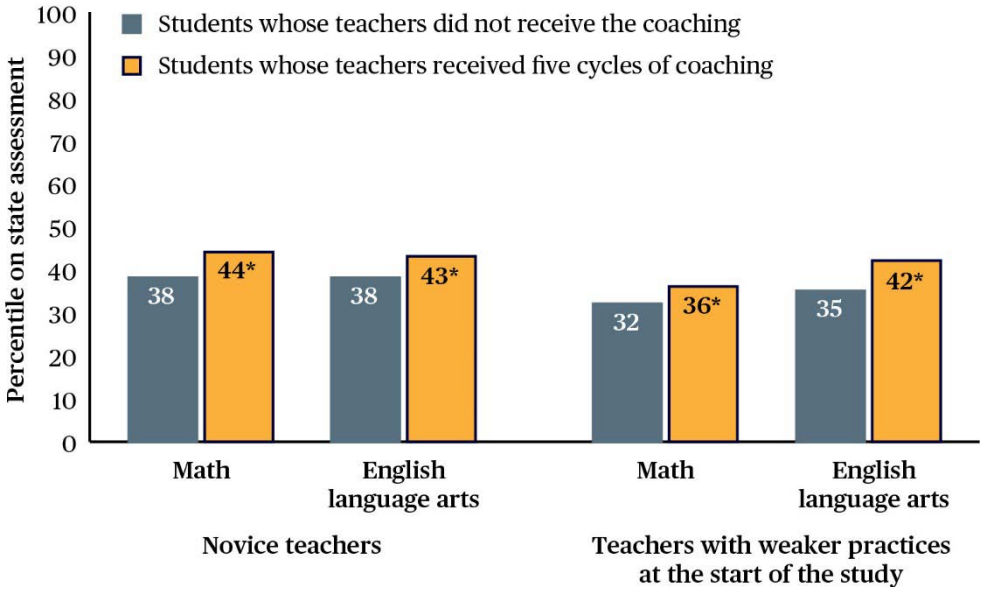
Source: Administrative student records for the 2017-2018 and 2018-2019 school years.

*Difference between five-cycle group and group that did not receive the coaching is statistically significant at the 0.05 level, two-tailed test.

- ***Five cycles of coaching improved achievement in both English language arts and math for students whose teachers might have had the greatest need for improvement.*** For novice teachers in their first five years of teaching, five cycles of coaching led to higher student achievement in both math and English language arts (Exhibit 4). For example, average English language arts scores for students of novice teachers in the five-cycle group were at the 43rd percentile, compared with average scores at the 38th percentile for students of novice teachers who did not receive the coaching. This is equal to a gain of about 2.5 months of learning. Similarly, among teachers with weaker classroom practices at the start of the study, five cycles of coaching led to higher student achievement in both subjects.[14] Studies suggest that novice teachers and those with weaker practices are more likely to work in schools with low test scores or students experiencing poverty, so improving these teachers' skills may be important to improve learning for students facing these challenges.[15]
- ***Five cycles of coaching was a cost-effective approach for improving student achievement.*** Five cycles of professional coaching based on videos improved student achievement at a cost of $228 per student on

average (see Appendix B, Section B.3.4). This is more cost-effective than other education strategies with evidence of effectiveness, including reducing class sizes, paying teachers extra for strong performance, and providing incentives for high-performing teachers to transfer to schools with low test scores (see Appendix B, Exhibit B.20). This suggests that five cycles of coaching is a cost-effective approach for improving student achievement.[16]

Exhibit 4. Effects of five cycles of coaching on student achievement for novice teachers and those with weaker practices at the start of the study



Source: Administrative student records for the 2017-2018 and 2018-2019 school years.

Note: Novice teachers are those who have been teaching for five years or less; teachers with weaker classroom practices at the start of the study are those who score in the bottom third of the sample on the measure of overall general teaching practices.

*Difference between five-cycle group and group that did not receive the coaching is statistically significant at the 0.05 level, two-tailed test.

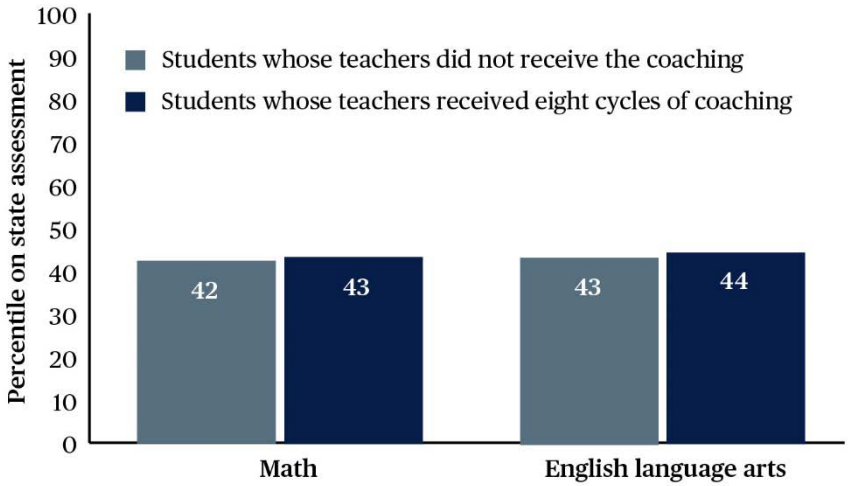
EIGHT CYCLES OF COACHING WAS NOT EFFECTIVE

Because policymakers have emphasized the importance of providing intensive, sustained professional development for teachers, the study also tested a version of the coaching that included an additional three cycles.[17] The eight-cycle version was designed to cover the same aspects of teaching and use the same approach as the five-cycle version. However, the additional cycles provided more opportunities for teachers to receive feedback and observe and reflect on their teaching. The coaches were expected to focus the additional cycles on practices that build students' understanding of the content being taught. For example, they might have focused on providing feedback that builds on students' responses, facilitating conversations that involve a majority of students, or helping students reflect on their thinking process. Given that teachers tend to struggle the most with these types of practices, the three additional cycles may lead to larger increases in student achievement. On the

other hand, a larger number of cycles would be more time-consuming for teachers and allow less time for them to practice and reflect between cycles, which could reduce the effect of the coaching on student achievement.

- ***Eight cycles of coaching did not improve student achievement.*** Students in the eight-cycle group had similar math and English language arts test scores as students whose teachers did not receive the coaching (Exhibit 5). In addition, eight cycles of coaching did not improve math or English language arts test scores among students of novice teachers or those with weaker classroom practices at the start of the study (Exhibit 6).[18] Eight cycles of coaching was less effective than five despite the fact that the same coaches delivered the coaching to both the eight and five cycle groups, and, as expected, teachers in the eight-cycle group received more coaching cycles on average than those in the five-cycle group.[19]

Exhibit 5. Effects of eight cycles of coaching on student achievement

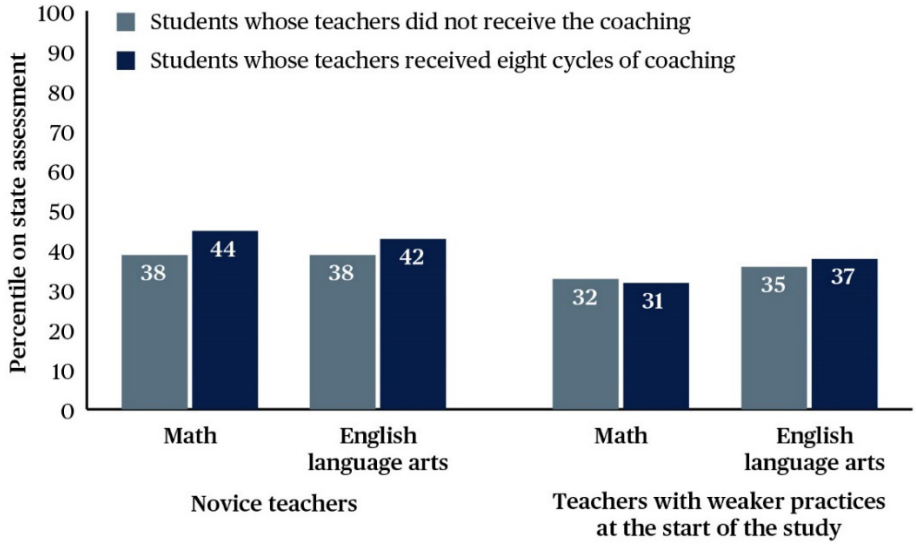


Source: Administrative student records for the 2017–2018 and 2018–2019 school years.

Note: Differences between eight-cycle group and the group that did not receive the coaching are not statistically significant at the 0.05 level, two-tailed test.

- ***Coaches adjusted the timing of the coaching to fit in eight cycles, but this may have limited its effect on student achievement.*** Teachers in the eight-cycle group spent more time in coaching activities during the school year than teachers in the five-cycle group. Across the school year, eight-cycle teachers spent about 30 more minutes watching videos of their teaching and 90 more minutes in coaching conferences compared to five-cycle teachers. However, coaching cycles for teachers in the eight-cycle group were shorter than those for the five-cycle group in order to fit the additional cycles into the school year. The average cycle lasted 28 days for teachers in the five-cycle group, compared to only 22 days for teachers in the eight-cycle group (see Appendix C, Exhibit C.14). In interviews conducted after the study ended, about half of the 15 coaches said they found it challenging to complete eight cycles in a single school year.[20] Although eight-cycle teachers spent more time in coaching activities overall, the shorter length of each cycle may have reduced the amount of time teachers had to work on individual practices and apply what they had learned in their classrooms before moving onto the next cycle.

Exhibit 6. Effects of eight cycles of coaching on student achievement for novice teachers and those with weaker practices at the start of the study



Source: Administrative student records for the 2017–2018 and 2018–2019 school years.

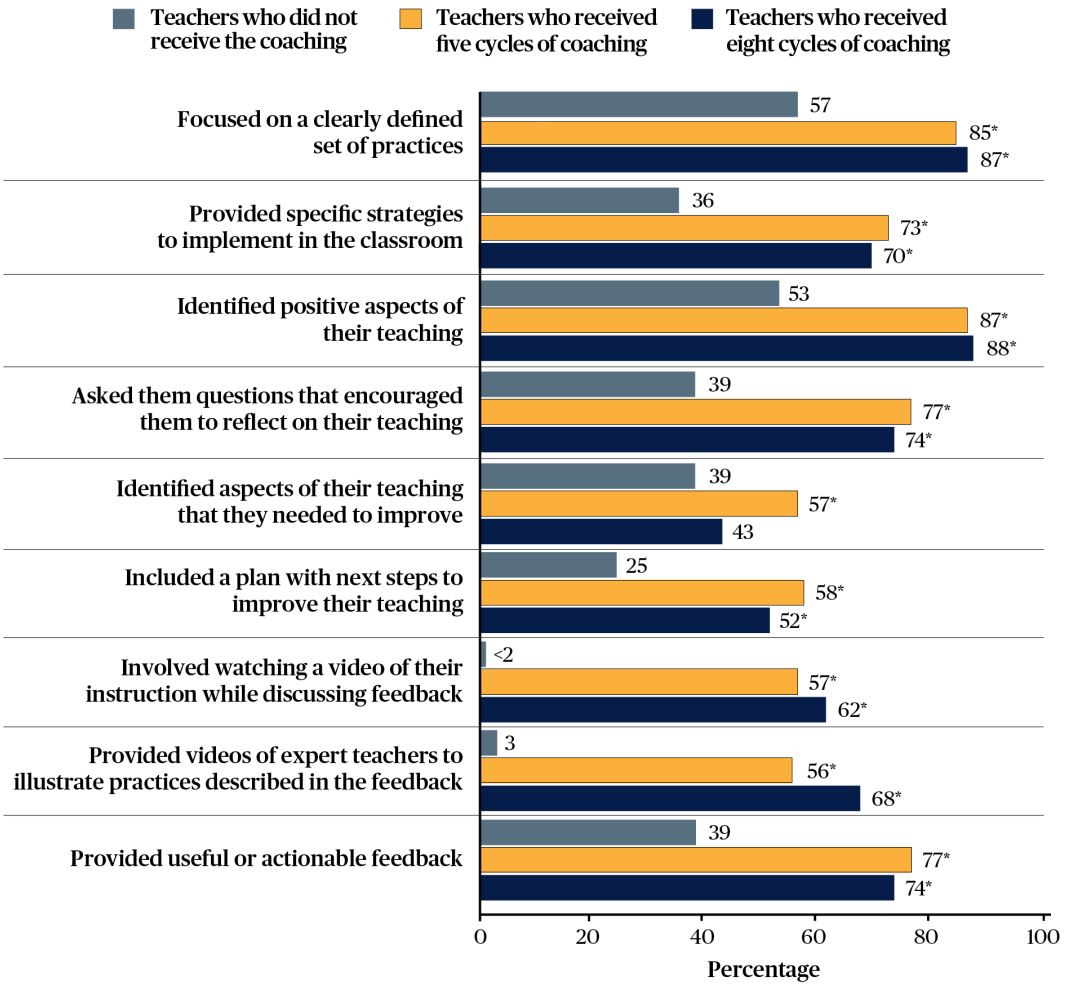
Note: Novice teachers are those who have been teaching for five years or less; teachers with weaker classroom practices at the start of the study are those who score in the bottom third of the sample on the measure of overall general teaching practices. Differences between eight-cycle group and the group that did not receive the coaching are not statistically significant at the 0.05 level, two-tailed test.

THE COACHING AFFECTED THE TYPE OF FEEDBACK TEACHERS RECEIVED ON THEIR TEACHING

The feedback provided through the study's coaching had features that may have differed from the informal feedback teachers often receive from principals or other teachers.[21] For example, prior to a coaching session, teachers were expected to reflect on their use of the targeted practices in the video clips using the written questions from the coach. In the coaching conference, coaches were expected to provide feedback to teachers, including detailed strategies for improving each of the targeted practices. For example, the teaching practice *facilitating conversations among students* included strategies such as teaching students to actively listen to other students; using a variety of different formats to promote student discussion; and asking open-ended questions to prompt discussion. These features of the coaching were designed to ensure that teachers received feedback tailored to their classrooms and focused on improving specific aspects of their teaching practices.

- ***As intended, the coaching increased the amount of feedback that teachers received.*** Teachers who received the study's coaching reported receiving more verbal feedback based on observations of their teaching than teachers who did not receive the coaching. On average, teachers in the five-cycle group received 222 minutes of verbal feedback based on observations of their teaching during the school year and teachers in the eight-cycle group received 255 minutes of verbal feedback, compared with 20 minutes for teachers who did not receive the coaching.
- ***The coaching changed the nature of the feedback teachers received in ways consistent with the program's design.*** For example, consistent with the coaching's focus on detailed aspects of teachers' practices, at least 70 percent of teachers in both coaching groups received feedback on specific strategies they could use in their classrooms, compared to 36 percent of teachers who did not receive the study's coaching (Exhibit 7). Teachers who received the coaching were also more likely to receive questions encouraging them to reflect on their own teaching (by at least 30 percentage points), and the feedback was more likely to refer to specific moments in their classroom observation (by at least 35 percentage points). [22, 23] This aligned with the written prompts that coaches provided with the video clips to help teachers reflect on specific moments or practices used in their teaching. The coaching also increased the proportion of teachers who received feedback on practices addressed by the coaching, such as leading discussions that build students' deeper understanding of the content being taught and supporting students' use of higher order thinking skills.[24] Although the study cannot determine whether these features of the coaching were responsible for its effects on student achievement, they are similar to the features of other coaching and feedback programs found to be effective in improving student achievement.[25]
- ***And teachers who received the coaching were more likely to report changing their teaching as a result of the feedback they received.*** Almost 90 percent of teachers in both coaching groups said they were more reflective about their teaching as a result of feedback they received, compared with only 57 percent of teachers who did not receive the coaching (Exhibit 8). Similarly, more than 80 percent of teachers in both coaching groups said they made a specific change to their teaching as a result of feedback they received, compared with only 51 percent of teachers who did not receive the coaching. In addition, more than 85 percent of teachers in both coaching groups thought the feedback was easy to understand, gave them specific ideas on how to improve, and would benefit students in the long run, whereas no more than 60 percent of teachers who did not receive the coaching shared these perceptions.

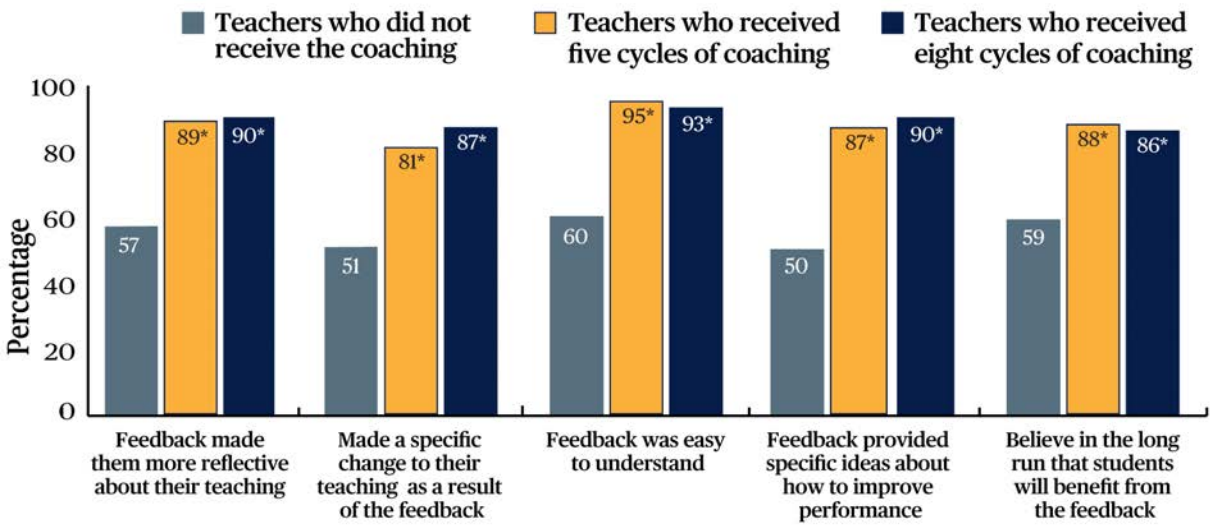
Exhibit 7. Differences in the type of feedback teachers received



Source: Teacher survey administered in spring 2019.

*Difference between five- or eight-cycle group and group that did not receive the coaching is statistically significant at the 0.05 level, two-tailed test.

Exhibit 8. Differences in teachers' perceptions of the feedback they received



Source: Teacher survey administered in spring 2019.

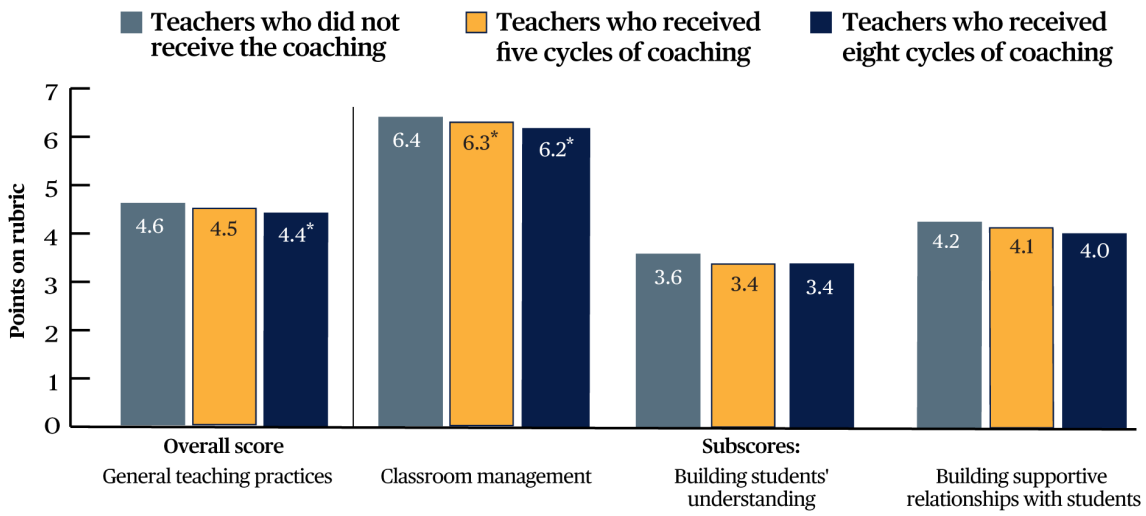
*Difference between five- or eight-cycle group and group that did not receive the coaching is statistically significant at the 0.05 level, two-tailed test.

- ***Although teachers described changing their practices, the study team did not observe improvements to teachers' practices, and found negative effects on some practices.*** Five cycles of coaching did not affect teachers' overall score on a classroom observation tool used to measure teachers' general teaching practices, and eight cycles lowered scores by 0.18 points on a 7-point scale (Exhibit 9).[26] For comparison, this negative effect is more than twice the gap in scores between novice and experienced teachers. The coaching had a negative effect on subscores measuring teachers' classroom management for both coaching groups, perhaps because the primary focus on building students' understanding of content diverted teachers' attention from classroom management. However, despite this negative effect, teachers in both coaching groups still scored very high on classroom management (an average score of at least 6 on a 7-point scale). The coaching did not affect subscores measuring practices related to building students' understanding of content even though this was its primary focus. Similarly, the coaching had no effect on subscores measuring practices related to building supportive relationships with students.

The analysis could not determine why teachers reported changing their practices but the study team did not observe improvements in the practices covered by the coaching. One possibility is that the classroom observation tool measured a broad range of practices, but the coaching may have only influenced a subset of practices that teachers covered during the coaching. For example, the observation tool measured 19 different practices related to building students' understanding of content. However, coaches focused on just a subset of these practices (on average, coaches covered 6 of the 19 practices with five-cycle teachers and 10 of the 19 practices with eight-cycle teachers). Another possibility is that the 30-minute observations used to measure teachers' practices did not provide sufficient opportunities to observe the specific practices teachers had improved. It is also possible that the study team would have observed improvements in teachers' practices if teachers had more time between coaching cycles to refine those practices or if the team's observations of

teachers' practices had occurred in the following school year, again after teachers had more time to refine their use of the practices.

Exhibit 9. Effects on teachers' classroom practice scores



Source: Teachers' classroom practices were measured based on video-recorded classroom observations from spring 2019. Overall scores on general teaching practices were measured on a 7-point scale using the Classroom Assessment Scoring System (CLASS).

*Difference between eight-cycle group and group that did not receive the coaching is statistically significant at the 0.05 level, two-tailed test.

Lessons Learned and Looking Forward

This study adds to a growing body of evidence on the effectiveness of teacher coaching by examining the effects of a specific type—coaching that provided feedback based on videos of teachers' instruction—and varying the amount of coaching provided. Although the study provides valuable evidence for districts, a few questions still need to be addressed.

- ***How important was the use of videos in the coaching?*** Although the study cannot isolate the effect of videos from the effects of other features of the coaching, videos of teachers' instruction played a central role in the coaching. Coaches selected short clips from the videos, used each clip for a specific purpose, and provided written questions to help teachers reflect on their clips. This study suggests that the use of video recordings can help teachers see and reflect on aspects of their teaching they may not otherwise notice, which may improve their teaching. Use of video recordings in coaching may have other potential benefits as well. It can provide more flexibility in when a coach or principal observes teachers in their classroom. It can also expand the pool of potential coaches to those located outside the district's geographic area. In addition, it can potentially reduce costs if teachers can easily record their own classrooms and coaches can more efficiently observe and provide feedback virtually than in person. However, more evidence is needed to understand how coaches can most effectively use videos to help teachers improve.
- ***How important was it that the coaching focused on general teaching practices rather than practices specific to a subject area like reading and math?*** Although the study cannot isolate the effect of providing feedback on general teaching practices from other aspects of the coaching, the findings are consistent with a growing number of studies that find this type of feedback has positive effects on student achievement.[27] These studies have used different approaches to providing feedback—including using peer teachers, formal performance evaluations, and professional coaches—and consistently find positive effects. This contrasts with the evidence on subject-specific coaching, which is more mixed. For example, several studies have found that literacy coaching is effective, but two studies of math-focused coaching did not find positive effects.[28] Although some teachers may benefit from subject-specific coaching, this study strengthens the evidence to support coaching teachers on their general teaching practices. Additionally, providing feedback on general practices may be more feasible and cost-effective for districts to implement—rather than having a separate coach for each subject area, a single coach provides feedback across subject areas.
- ***What are the long-term effects of the coaching?*** Prior studies suggest that coaching may have effects on student achievement in the school year *after* teachers are coached.[29] Coaching may have long-term effects if teachers sustain improvements to their practices or continue to improve their skill with these new practices over time. Unfortunately, the COVID pandemic led to the cancellation of state assessments in spring 2020, which prevented this study from measuring effects of the coaching in the year after teachers received it. Thus, it is unclear whether the effects of five cycles of coaching would grow or be sustained the year after the coaching was delivered. It is also unclear whether eight cycles of coaching would improve student achievement in a subsequent year.

The study suggests that individualized, video-based coaching for teachers focused on general teaching practices can be a cost-effective approach for improving student achievement. Teacher coaching represents a potentially important approach for districts and schools seeking to improve teacher effectiveness—including novice teachers and teachers with weaker practices. Building evidence to address these remaining questions is critical for ensuring that educators and policymakers design coaching in a way that best supports teachers and promotes student learning.

ENDNOTES

[1] Chetty et al. 2014; Jackson 2018.

[2] TNTP 2015; U.S. Department of Education 2016a, 2016b.

[3] A 2009 review of existing research proposed five core features of effective professional development: that it was of sufficient duration; focused on subject matter content; involved groups of teachers from the same school, grade, or department; provided opportunities for teachers to engage in active learning; and was consistent with school, district, and state reforms and policies (Desimone 2009). However, three large-scale random assignment studies that evaluated professional development with these features found that it improved teachers' knowledge and classroom practices but not student achievement (Garet et al. 2016a). In addition, although recent meta-analyses find positive effects of professional development on student achievement, they do not identify a consistent set of features that make professional development effective (Kennedy 2016; Kraft et al. 2018; Lynch et al. 2019; Didion et al. 2020). For example, four of these meta-analyses found that the duration or intensity of professional development is not related to its effectiveness (Kennedy 2016; Kraft et al. 2018; Lynch et al. 2019; Didion et al. 2020).

[4] Garet et al. 2017; Allen et al. 2011, 2015; Taylor and Tyler 2012; Steinberg and Sartain 2015; Papay et al. 2020.

[5] For example, a recent meta-analysis found that the amount of coaching provided to teachers was not related to its effects on student achievement (Kraft et al. 2018). Also, studies find that teachers' overall teaching practices are related to student achievement, yet these studies have not identified which specific practices are most important for improving student achievement (Kane and Staiger 2012; Garrett and Steinberg 2015; Chaplin et al. 2014).

[6] Allen et al. 2011, 2015.

[7] Kane et al. 2015; Roth et al. 2011; Sherin and van Es 2005, 2009.

[8] Districts would need to weigh the benefit of relying on a single coach across subject areas against the potential benefits of having subject matter experts provide coaching in specific subject areas. Kraft et al. (2018) found that coaching focused on a specific subject area and coaching focused on general teaching practices had similar effects on student achievement (0.51 standard deviations for subject-specific coaching and 0.47 standard deviations for coaching on general teaching practices).

[9] The coaching focused on a set of 40 teaching practices from the Classroom Assessment Scoring System (CLASS) classroom observation tool. Appendix A describes the CLASS and lists the practices that it covers.

[10] For example, providing teachers four rounds of feedback based on classroom observations has improved student achievement in English language arts (Garet et al. 2017) and math (Taylor and Tyler 2012).

[11] The study had planned to measure the effects of the coaching on student achievement in the year teachers received it and in the year after they received it. However, the study was not able to measure effects of the coaching in the year after teachers received it because COVID-19 led to the cancellation of state assessments.

[12] Appendix B provides more information on the estimation of the effects of the coaching, the conversion to percentile units, and the conversion to months of learning.

[13] Average math scores for students of teachers in the five-cycle group were at the 45th percentile, compared with average scores at the 42nd percentile for students whose teachers did not receive coaching. However, the estimated effect on student math score was not statistically significant at the 5 percent level, with a *p*-value of 0.07.

[14] Teachers with weaker teaching practices at the start of the study are those who scored in the bottom third of the sample on the CLASS. Scores were based on observations of their general teaching practices at the start of the study school year.

[15] Isenberg et al. 2016; Center for Education Policy Research 2012.
[16] Appendix B provides information about how the cost effectiveness of the coaching was measured and compares the cost effectiveness of the coaching to other strategies to improve student achievement. The comparisons focused on other strategies that, like the study's coaching, (1) seek to improve student achievement by influencing teachers' effectiveness, (2) could plausibly be implemented in grades 4 and 5, and (3) have rigorous evidence of effectiveness and detailed information on costs from existing studies.
[17] The Every Student Succeeds Act (2015) provides federal funding for professional development activities that are "sustained (not stand-alone, one-day, or short-term workshops), intensive, collaborative, job-embedded, data-driven, and classroom-focused."
[18] Although average scores in both math and English language arts were higher for students taught by novice teachers in the eight-cycle group than for students taught by teachers who did not receive the coaching, the study could not definitively conclude that eight cycles of coaching improved achievement for students of novice teachers. The estimated effects on student math and English language arts scores were not statistically significant at the 5 percent level, with *p*-values of 0.08 and 0.11, respectively (see Appendix C, Exhibit C.2).
[19] Five-cycle teachers completed 4.3 coaching cycles on average and eight-cycle teachers completed 7.0 coaching cycles on average (see Appendix A, Exhibit A.6).
[20] During semi-structured interviews with coaches, the coaches were asked, "To what extent do you view the following as challenges of having 8 cycles of coaching instead of 5 cycles, using the scale (1) not at all, (2) to a small extent, (3) to a moderate extent, or (4) to a great extent?" In response to the challenge "Fitting in all 8 cycles before the end of the school year," 7 of the 13 coaches who participated in the interviews indicated the challenge affected them to a moderate or great extent.
[21] A nationally representative survey found that 71 percent of teachers receive feedback based on informal observations by principals and 54 percent receive this type of feedback from other teachers (Tuma et al. 2018).
[22] The percentage of teachers reporting that the feedback they received provided questions that encouraged them to reflect on their own teaching was 77 percent for the five-cycle group, 74 percent for the eight-cycle group, and 39 percent for teachers who did not receive the coaching (Appendix C, Exhibit C.4). Similarly, the percentage of teachers reporting that the feedback they received provided specific techniques or strategies they could implement in their classroom was 73 percent for the five-cycle group, 70 percent for the eight-cycle group, and 36 percent for teachers who did not receive the coaching (Appendix C, Exhibit C.4).
[23] Five cycles of coaching increased the proportion of teachers who received feedback that identified aspects of their teaching that they needed to improve, but eight cycles of coaching did not. However, this finding does not suggest that this particular feature of the feedback, rather than the difference in number of coaching cycles, explained the differences in coaching effectiveness for the five- and eight-cycle groups. The same coaches delivered feedback to teachers in both groups, and there were no systematic differences in the types of feedback they were expected to provide to the two groups. Because the specific features of the coaching were not randomly assigned but were determined based on coaches' discretion and teachers' needs, the study cannot conclude that differences in these features across the five- and eight-cycle groups led to differences in effects on student achievement.
[24] For example, about three-quarters of teachers in the coaching groups reported receiving feedback on leading discussions that build students' deeper understanding of the content being taught compared with 33 percent of teachers who did not receive the coaching (see Appendix C, Exhibit C.6).
[25] For example, Garet et al. (2017) found that providing teachers individualized feedback on a clear set of teaching practices in a classroom observation rubric led to positive effects on students' math achievement.

[26] The study measured teachers' general classroom practices using the CLASS. Scores were based on video-recorded classroom observations from the spring of the study school year. Because the coaching improved student achievement in English language arts, the study also examined whether the coaching affected teaching practices specific to English language arts, as measured by a well-established classroom observation tool focused on these practices (the Protocol for Language Arts Teaching Observations). However, the coaching did not affect teachers' overall scores on these English language arts-focused practices (see Appendix C, Exhibit C.11).

[27] Garet et al. 2017; Allen et al. 2011, 2015; Taylor and Tyler 2012; Steinberg and Sartain 2015; Papay et al. 2020.

[28] Kraft et al. (2018) conducted a meta-analysis that included 34 studies of literacy coaching. Two studies examined math coaching and neither found an impact on student achievement. Garet et al. (2016b) studied a professional development program that included feedback from professional coaches on teachers' math instruction. Kraft and Hill (2018) also studied feedback from professional coaches based on videos of teachers' math instruction.

[29] Allen et al. 2011, 2015; Taylor and Tyler 2012.

REFERENCES

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

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

Center for Education Policy Research. "Do Low-Performing Students Get Placed With Novice Teachers?" Cambridge, MA: Harvard University, Center for Education Policy Research, 2012.

Chaplin, Duncan, Brian Gill, Allison Thomkins, and Hannah Miller. "Professional Practice, Student Surveys, and Value-Added: Multiple Measures of Teacher Effectiveness in the Pittsburgh Public Schools." Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, July 2014.

Chetty, Raj, John Friedman, and Jonah Rockoff. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." *American Economic Review*, vol. 104, no. 9, 2014, pp. 2633–2679.

Desimone, Laura. "Improving Impact Studies of Teachers' Professional Development: Toward Better Conceptualizations and Measures." *Educational Researcher*, vol. 38, no. 3, 2009, pp. 181–199.

Didion, Lisa, Jessica Toste, and Marissa Filderman. "Teacher Professional Development and Student Reading Achievement: A Meta-Analytic Review of the Effects." *Journal of Research on Educational Effectiveness*, vol. 13, no. 1, 2020, pp. 29–66.

Every Student Succeeds Act, 114-95 U.S.C. § 1601 (2015).

Garet, Michael, Jessica Heppen, Kirk Walters, Julia Parkinson, Toni Smith, Mengli Song, Rachel Garrett, Rui Yang, and Geoffrey Borman. "Does Content-Focused Teacher Professional Development Work? Findings from Three Institute of Education Sciences Studies." Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2016a.

Garet, Michael, Jessica Heppen, Kirk Walters, Julia Parkinson, Toni Smith, Mengli Song, Rachel Garrett, Rui Yang, Geoffrey Borman, and Thomas Wei. "Focusing on Mathematical Knowledge: The Impact of Content-Intensive Teacher Professional Development." NCEE 2016-4010. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2016b.

Garet, Michael, Andrew Wayne, Seth Brown, Jordan Rickles, Mengli Song, and David Manzeske. "The Impact of Providing Performance Feedback to Teachers and Principals." NCEE 2018-4001. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, December 2017.

Garrett, Rachel, and Matthew Steinberg. "Examining Teacher Effectiveness Using Classroom Observation Scores: Evidence from the Randomization of Teachers to Students." *Educational Evaluation and Policy Analysis*, vol. 37, no. 2, 2015, pp. 224–242.

Kane, Thomas, Hunter Gehlbach, Miriam Greenberg, David Quinn, and Daniel Thal. "The Best Foot Forward Project: Substituting Teacher-Collected Video for In-Person Classroom Observations. First Year Implementation Report." Cambridge, MA: Harvard University, 2015.

Isenberg, Eric, Jeffrey Max, Phillip Gleason, Matthew Johnson, Jonah Deutsch, and Michael Hansen. "Do Low-Income Students Have Equal Access to Effective Teachers? Evidence from 26 Districts." NCEE 2017-4007. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2016.

Jackson, Kirabo. "What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes." *Journal of Political Economy*, vol. 126, no. 5, 2018, pp. 2072–2107.

Kane, Thomas, and Douglas Staiger. "Gathering Feedback for Teaching: Combining High-Quality Observations with Student Surveys and Achievement Gains." Seattle, WA: Bill & Melinda Gates Foundation, 2012.

Kennedy, Mary. "How Does Professional Development Improve Teaching?" *Review of Educational Research*, vol. 86, no. 4, 2016, pp. 1–36.

Kraft, Matthew, David Blazar, and Dylan Hogan. "The Effect of Teacher Coaching on Instruction and Achievement: A Meta-Analysis of the Causal Evidence." Review of Educational Research, vol. 88, no. 4, 2018, pp. 547–588.

Kraft, Matthew, and Heather Hill. "Developing Ambitious Mathematics Instruction Through Web-Based Coaching: A Randomized Field Trial." *Brown University Working Paper*. Providence, RI: Brown University, 2018.

Lynch, Kathleen, Heather Hill, Kathryn Gonzalez, and Cynthia Pollard. "Strengthening the Research Base that Informs STEM Instructional Improvement Efforts: A Meta-Analysis." *Educational Evaluation and Policy Analysis*, vol. 41, no. 3, 2019, pp. 260–293.

Papay, John, Eric Taylor, John Tyler, and Mary Laski. "Learning Job Skills from Colleagues at Work: Evidence from a Field Experiment Using Teacher Performance Data." *American Economic Journal: Economic Policy*, vol. 12, no. 1, 2020, pp. 359–88.

Roth, Kathleen, Helen Garner, Catherine Chen, Meike Lemmens, Kathleen Schwille, and Nicole Wickler. "Video-Based Lesson Analysis: Effective Science PD for Teacher and Student Learning." *Journal of Research in Science Teaching*, vol. 48, 2011, pp. 117–148.

Sherin, Miriam Gamoran, and Elizabeth A. van Es. "Using Video to Support Teachers' Ability to Notice Classroom Interactions." *Journal of Technology and Teacher Education*, vol. 13, no. 3, 2005, pp. 475–491.

Sherin, Miriam Gamoran, and Elizabeth A. van Es. "Effects of Video Club Participation on Teachers' Professional Vision." *Journal of Teacher Education*, vol. 60, no. 1, 2009, pp. 20-37.

Steinberg, Matthew, and Lauren Sartain. "Does Teacher Evaluation Improve School Performance? Experimental Evidence from Chicago's Excellence in Teaching Project." *Education Finance and Policy*, vol. 10, no. 4, 2015, pp. 535–572.

Taylor, Eric, and John Tyler. "The Effect of Evaluation on Teacher Performance." *The American Economic Review*, vol. 102, no. 7, 2012, pp. 3628–3651.

TNTP. "The Mirage: Confronting the Hard Truths About Our Quest for Teacher Development." New York City, NY: TNTP, 2015.

Tuma, Andrea Prado, Laura Hamilton, and Tiffany Tsai. "A Nationwide Look at Teacher Perceptions of Feedback and Evaluation Systems: Findings from the American Teacher Panel." Santa Monica, CA: RAND, 2018.

U.S. Department of Education. "Findings from the 2015-2016 Survey on the Use of Funds Under Title II, Part A: Subgrants to LEAs." Washington, DC: U.S. Department of Education, August 2016a.

U.S. Department of Education. "Findings from the 2015-2016 Survey on the Use of Funds Under Title II, Part A: Subgrants to State Activities Funds." Washington, DC: U.S. Department of Education, 2016b.

ACKNOWLEDGMENTS

We are grateful to the many people who contributed to this study and report.

First and foremost, we thank the teachers who participated in the study, along with the other school and district staff who supported the study and assisted with data collection.

We thank the leaders and coaches from Teachstone who implemented the study's coaching, including Vicki Kintner-Duffy, Mamie Morrow, Anna Savitsky, Sarah Hadden, Eydie Anchondo, Nicole Davis, Bria Gardner, Rosemary Molina, Kristin Valdes, Alka Franceschi, Rootie Taylor, Aisha Loucas, Kelsey Trudo, Stephanie Nichols, Joe Pierce, Lisa Wierck, Destiny Woodbury, Leah Osbourne, Michelle Schmidt, Marcella Corrales, and Vivian Polo. We also thank the team of staff from WestEd that supported implementation of the coaching. They included Natalie Lacireno-Paquet, Edward Britton, Betsy McCarthy, Lauren Agnew, Candice Bocala, Cindy Caldwell, Rebeca Diaz, and Claire Morgan.

A large team of staff from Mathematica, and Decision Information Resources helped recruit schools into the study. They included Tim Silva, Eric Zeidman, Steve Bates, Sara Woodruff, Moira McCullough, Megan Davis, Kevin Kelly, Dan Friend, Charlotte Cabili, Carol Pistorino, Brittany English, Anne Scherer, Amanda Lee, Virginia Knechtel, Steven Malick, Scott Peecksen, Ronald McCowan, Libby Makowsky, Jeffrey Max, Jacob Hartog, Frances Bailey, Emilyn Whitesell, and Antwanette Hill.

Many staff helped collect and analyze data for the study. Sheila Heaviside led the administration of teacher surveys, with support from Tim Bruursema, Ann Ponti, Tong Li, John McCarthy, and Lucy Tindall. Tim Bruursema led the collection of video-recorded classroom observations, with support from Maya Reid, Eliza Abendroth, Alex Hollister, Doug Ortiz, Sarah Wagner, Eric Brower, Daniel Welsh, Richard Godwin, Joan Gutierrez, Rhonda Lofton, and Evelyn Schlinski. IRIS Connect provided the camera equipment used to video record teachers' classrooms, and TeachForward provided the online platform used to store and code the classroom videos. Carol Pistorino and Cynthia Walsh from DIR oversaw the coding of the videos. A team of staff from Mathematica, led by Megan Davis and Emma Ernst, collected administrative data for the study. They included Emily Rosen, Tori Rockwell, Jessie Mazeika, and Brittney Gionfriddo. Maria Bartlett, Kerry Schellenberger, Aldo Iturrios, Alma Vigil, and Morgan Kolarich conducted the analysis, with guidance from Emma Ernst.

The study received useful advice from a technical working group, convened by the research team over the course of the study. The group included David Chard, Jan Dole, Jim Kemple, Judith Warren Little, Jim Wyckoff, Pam Grossman, Russell Gersten, Steve Rivkin, Sharon Vaughn, and Tonya Wolford.

Suzanne Wilson of the University of Connecticut helped shape and guide the evaluation and provided expert contributions and advice at various stages. Mark Dynarski of Pemberton Research provided important input on the study design. Catherine McClellan from Clowder Consulting provided feedback on the coding of classroom observations. Alison Wellington provided expert reviews of the report.

Jennifer Brown edited the report; Gwyneth Olsen, Yvonne Marki, and Michael King created graphics; and Sharon Clark prepared the report for publication.

DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST

The research team for this evaluation included staff from Mathematica and its subcontractors and consultants, WestEd, Decision Information Resources, Pemberton Research, Clowder Consulting, TeachForward, IRIS Connect, and Suzanne Wilson of the University of Connecticut. None of the research team members has financial interests that could be affected by findings from this evaluation. Similarly, none of the participants in the study's technical working group meetings has financial interests that could be affected by findings from the evaluation.