Evaluation of the Teacher Potential Project

July 30, 2019

Sarah Dolfin, Scott Richman, Jane Choi, Andrei Streke, Cheryl DeSaw, Alicia Demers, Dmitriy Poznyak

Submitted to:

EL Education
247 West 35th Street #800
New York, NY 10001
Beth Miller, Contract Officer

Submitted by:

Mathematica
1100 First Street, NE
12th Floor
Washington, DC 20002
(202) 484-9220 Phone
(202) 863-1763 Fax
Contract Number: 40342-C-140421

This page has been left blank for double-sided copying.

Acknowledgments

This study was made possible by the efforts of many individuals. We thank Meg Riordan, our project officer, for supporting our work; Beth Miller, for her feedback, information, and support throughout the project; and Caitlin McKenzie, Christina Lesh, and others at EL Education who helped recruit schools and provided important information about the implementation of TPP. We also appreciate valuable feedback from others at EL including Scott Hartl.

This report would not have been possible without contributions from many other people at Mathematica. Recruitment efforts were supported by Jesse Chandler, Leonard Brown, Anna Situ, and Roland Scurato. Random assignment was supported in the planning phase by Carol Razafindrakoto and was conducted by Molly Crofton. Classroom observations were supported by Nancy Collins, Markelia Cooper, Elizabeth Dawes, Mary Diffley, Caileen Geary, Emily Geiger-Medina, Richard Godwin, Susan Golden, Lauren Harris, Emma Klimek, Robin Lewis, Karen Markowski, Kayley Pettoruto, Annastas Poplus, Shamama Siddiqui, Kimberly Smith, Kathleen Spitz, and Pamela Zendt. The student literacy task design and grading was supported by Meg Bishop, Bridget Kelly, Grace Lim, Brianna Noll, Andrea Parker, and Sara Woodruff. Survey design and processing were supported by Gavin Fugere, Dan O'Connor, and Kester St. Kitts. Programming for the study was led by Emma Ernst and conducted by Kathryn Cronquist, Micah Melia, Kelley Monzella, and Hanzhi Zhou. Armando Yañez assisted with identifying and reviewing related literature. Beau Smit, Max Pohlman, and Danni Tu supported the inter-rater reliability analyses. Zachary James, Frank Masters, and Moriah West provided assistance with project management. Roberto Agodini, Claudia Gentile, and Eric Isenberg provided valuable leadership earlier in the project. Phil Gleason, Joshua Haimson, and Peter Schochet provided critical technical review and feedback. Sharon Clark and Cindy Castro led the production of the report, which was edited by Kerry Kern.

Finally, we sincerely appreciate the willingness of principals and teachers to participate in the evaluation, and are grateful to district and school staff for providing administrative data.

This report was developed under a validation grant (U411B130041) awarded to EL Education by the U.S. Department of Education through the Investing in Innovation (i3) program. The opinions, findings, conclusions, and recommendations contained in this report do not necessarily reflect the view of the U.S. Department of Education or other funders.

This page has been left blank for double-sided copying.

CONTENTS

TABLES

FIGURES

This page has been left blank for double-sided copying.

EXECUTIVE SUMMARY

There is a growing need to help students develop the skills required to become college- and career-ready, as defined by rigorous state standards that have been implemented by a vast majority of states, such as the Common Core State Standards (CCSS). It is critical that teachers have the supports needed to provide effective instruction in these areas; this is especially true for novice teachers, who face the dual challenge of becoming effective teachers and meeting the new standards.

Integrating teacher professional learning with educative curriculum has emerged as a promising way to both build teacher capacity and support student achievement. In particular, this type of approach helps make professional development supports more relevant for teachers to apply in their classrooms because it is aligned to the curriculum, which could potentially address the issues that much of the current professional development does not meet teachers' needs (Bill & Melinda Gates Foundation 2014; Calvert 2016) and may not be associated with improvement in teacher performance (TNTP 2015).

EL Education designed the Teacher Potential Project (TPP) to build the instructional capacity of English language arts (ELA) teachers, and novice ELA teachers in particular. TPP aims to enhance ELA teachers' instructional practices in areas aligned to CCSS and to foster student learning and achievement outcomes. TPP provides an ELA curriculum paired with teacher professional learning supports to help teachers deliver high-quality ELA instruction. The open-source, standards-aligned curriculum for grades 3 through 8 is aligned to CCSS. The teacher professional development learning supports include (1) on-site institutes; (2) on-site and video-based coaching from EL Education coaches, including observations, modeling, lesson studies, and fostering professional learning communities within the school; and (3) access to online supports, including a range of online professional development materials and opportunities to participate in online communities of practice and interactive webinars.

In 2013, EL Education was awarded a five-year, $11.9 million Investing in Innovation validation grant by the U.S. Department of Education. The purpose of the grant was to help build upon the existing evidence for EL Education's curriculum and professional development model by assessing the effectiveness of TPP. Mathematica, as the third-party evaluator for the grant, designed and conducted the Evaluation of the Teacher Potential Project, which includes an implementation evaluation and a teacher and student impact evaluation. The study of TPP makes several important contributions to the literature evaluating paired curriculum and PD programs: it uses rigorous group designs, evaluates the impact of one and two years of program implementation, and examines broad outcomes on both teacher instructional practice and student ELA achievement. This report describes the study and its findings.

A. Study design and data collection

The evaluation uses a randomized controlled trial (RCT) design to assess the impacts of a single year of TPP implementation on the instructional practice outcomes of teachers and achievement

outcomes of students in the study during that year. We also use this design to assess impacts on student achievement among all schools in the study—those that engaged in one year of TPP implementation and others that engaged in two years of implementation.

The study team recruited 19 relatively high-need school districts across the United States in three cohorts that participated during the 2014–2015, 2015–2016, and 2016–2017 school years. Within each cohort, participating districts selected elementary and middle schools meeting the eligibility criteria to participate in the study, for a total of 79 schools. Schools were eligible if they had at least one novice (defined in this study as those in those with zero to three years of full-time teaching experience) ELA teacher and at least 45 minutes of ELA classes, among other criteria. The study team randomly assigned schools to treatment and control conditions within matched pairs of schools within districts.[1] Treatment schools participated in TPP for one school year, which involved their ELA teachers using TPP's curriculum and receiving TPP's professional development (PD) supports, while control schools and their ELA teachers continued to provide their typical ELA curricula and teacher supports. Of the 79 schools, 40 were assigned to the treatment condition and 39 were assigned to the control condition; 70 schools (35 treatment and 35 control) in 18 districts went on to participate in the study. The participating schools had a relatively lower socioeconomic status (71 percent of students were eligible for free or reduced-price lunch), served a largely minority race and ethnicity population of students (46 percent were black non-Hispanic, 21 percent were Hispanic), and had below-average student achievement scores. There were no statistically significant differences in these features between treatment and control schools.

In addition to the RCT, the study team designed a two-year quasi-experimental design (QED) study to assess the impact of extending implementation of TPP to a second year. EL Education recruited 22 of the study schools (10 treatment, 12 control) in five districts in Cohort 3 to participate in a second year of the study in the following 2017–2018 school year. Treatment schools that chose to participate for a second year would continue to implement TPP for a second year, while control schools continued to provide their typical ELA curricula and teacher supports. The schools participating in the two-year QED study had a somewhat lower socioeconomic status (61 percent of students were eligible for free or reduced-price lunch) and had slightly higher baseline ELA achievement than the full sample of schools participating in the RCT. There were no statistically significant differences in these features between treatment and control schools.

The study team collected a variety of data for the evaluation of TPP. We collected rosters of teachers in study schools to identify the population of teachers for the study. Teacher surveys were administered and classroom observations were conducted in fall and spring each year the schools participated to gather information on instructional practices for the teacher impact evaluation. In addition to the teacher surveys, teacher professional development exit surveys and TPP coach surveys were administered to gather information on teachers' perceptions of the TPP

[1] If a state had two participating districts that each identified one school for inclusion in the study, we randomly assigned one district to the treatment condition and one to the control condition. This occurred in two states, one in Cohort 2 and one in Cohort 3.

institutes and coaches' perceptions of TPP implementation and teacher participation for the implementation evaluation. Student administrative records were collected from districts to obtain information on student ELA test scores before and after the implementation of TPP, and a literacy task was administered to students of Cohort 3 novice teachers in spring 2017, for the student impact evaluation.

B. Key implementation findings

The implementation evaluation is designed to address the research question: "Is there evidence of intervention fidelity?" The implementation evaluation supports replication and scaling of TPP. Its findings indicate the extent to which treatment schools took up and engaged in the services that TPP offered, which can help with the interpretation of the impact study findings.

The implementation evaluation consists of an analysis that identifies the core components of the program, defines the necessary implementation thresholds for these core components, and assesses the extent to which TPP was successfully implemented in treatment schools. These analyses use data collected through coach and teacher reports and classroom observations.

1. The TPP ELA curriculum was implemented in all schools

The TPP CCSS-aligned ELA curriculum was implemented in all treatment study schools that participated in the first and second year of TPP. In their surveys, TPP coaches described successes with the curriculum implementation, which included teachers following the lesson plans included in the curriculum modules. However, teachers also experienced issues with delivering the curriculum modules, such as not teaching all the aspects of a module, as they were learning and acquainting themselves with the curriculum.

2. There was high school-level implementation fidelity of the TPP professional development components in the first and second years of TPP among the novice ELA teachers

For each of the three PD components—institutes, coaching, and online supports—we created teacher-level participation scores based on information provided by coaches' reports and teachers' self-report and using these, we developed school-level fidelity scores to gauge whether each component was implemented with fidelity. If more than 50 percent of the schools had high fidelity with a PD component, that component was deemed as being implemented with fidelity. Separate analyses were performed when only including novice teachers in the sample of teachers at a given school and when including all ELA teachers (novices and experienced), along with assessing implementation among schools receiving one year and two years of TPP.

Overall, the institutes component of TPP's PD was implemented with fidelity with schools in their first and second year of TPP among both novice and all teachers in the study. Teachers and coaches both reported that the institutes were useful in helping teachers to implement the ELA curriculum, although coaches reported that scheduling and teachers' availability for these institutes was a common challenge. The TPP coaching component was also implemented with

fidelity across study years and among novice and all teachers. While TPP coaches felt that teachers were receptive to the feedback and coaching, teacher availability for coaching during the school day sometimes posed a challenge for providing PD support. Finally, while there was implementation fidelity of the online supports in schools in their first year among novice ELA teachers, this was not the case with novice ELA teachers in their second year of TPP nor among all ELA teachers during either year of TPP implementation.

C. Key impact findings

The impact evaluation is designed to address two general research questions about impacts of TPP on teachers and students. First, it answers the question, "What is the impact of TPP on the ELA instructional practices of novice teachers (defined in this study as those in those with zero to three years of full-time teaching experience) and all teachers in upper elementary and middle school grades?" Second, it answers the question, "What is the impact of TPP on the ELA achievement of upper elementary and middle school students in the classrooms of novice and all teachers?"

1. TPP had positive impacts on teacher instructional practices

We examined impacts for teachers who experienced one or two years of TPP supports to understand how these different durations of engagement might change teachers' instructional practices. We examined these impacts separately for novice ELA teachers as well as all ELA teachers to learn how TPP supports affect not just teachers in general, but also teachers who are relatively new to the profession specifically. In order to define complex teacher practice outcomes in a robust way, we combined information from specific teacher survey and classroom observation items related to instructional practice into 31 constructs within 16 topics related to the areas of general instruction, reading and writing instruction, and classroom management and environment. The general instruction area captures instructional practices that support student learning, which included teachers engaging in the following practices: having well-structured lessons; connecting students' learning to their prior knowledge, personal lives, or the real world; supporting students' higher-order thinking and content knowledge development; encouraging students' participation in class and discussions; and supporting students' responsibility for their own learning. The reading and writing instruction area captures instructional practices that support students' engagement with texts and reading and writing practices that are the foci of CCSS, which included integrating academic vocabulary; having students engage in multiple types of writing and write for multiple purposes; having students engage with texts by reading, writing, and/or speaking about texts; using evidence from text to support their ideas; and focusing on developing students' writing conventions. The classroom management and environment area captures instructional practices that support classroom management and create a positive classroom environment. Impacts were estimated through regression analysis that compared practice outcomes of treatment teachers to those of control teachers, controlling for their students' reading and math pre-test scores and background characteristics, and allowing for the clustering of teachers within schools.

The study found that one year of TPP had statistically significant, positive impacts on all treatment teachers' overall ELA instructional practices, which included practices in the general instruction, reading and writing, and classroom management and environment areas. There were also significant, positive impacts for particular ELA-specific instructional practices, such as engaging students in reading, writing and/or speaking about texts; and supporting students' use of text evidence, and general instructional practices, which included providing students with connected lessons and supporting their higher order thinking. Among teachers who received a second year of TPP, this study did not find statistically significant impacts, likely due to the small sample size of teachers in the analyses. However, earlier research from this study found statistically significant impacts on the specific instructional practices of more often encouraging students' higher-order thinking, asking students to use evidence from texts, and engaging students in reading, writing, and/or speaking about texts among teachers who received a second year of TPP (Choi et al. 2018).

2. TPP had positive impacts on student ELA achievement

We examined a variety of impacts to understand how exposure to different durations and types of engagement with TPP affects students. Impacts were estimated through regression analysis that compared scores of students of treatment teachers to those of students of control teachers on state-administered ELA assessments that were standardized using means and standard deviations of a national norming population. The analyses controlled for student reading and math pre-test scores; student, teacher, and school background characteristics; and year, district, and grade level; and allowed for the clustering of students within schools.

The study found no effects on students' achievement at the end of the first year that their teachers engaged with TPP. However, in the year after teachers engaged in a single year of TPP, there were positive impacts on achievement approaching the 5 percent level of significance for students in their classrooms. This impact had an effect size of 0.06 standard deviations.

There was a positive and significant impact on the achievement of students in the classrooms of teachers in their second year of engagement with TPP. This impact had an effect size of 0.10 standard deviations, which can be interpreted as roughly 1.4 months of typical student improvement, or moving an average student scoring at the 50th percentile to the 54th percentile. These impacts were confined to the students of teachers who engaged with TPP directly and did not appear to affect students of other nonstudy ELA teachers in the study schools.

Among all schools that engaged with TPP—for a single year or for two years of implementation—there was a positive and significant impact on student achievement in the second year, with an effect size of 0.08 standard deviations. There were no significant impacts on the achievement of students in the classrooms of novice teachers during their first or second year of engagement with TPP.

The study also estimated impacts on students' scores on an opinion or argument writing task at the end of their teachers' first year of engagement with TPP. Impacts were estimated for a subset of Cohort 3 novice ELA teachers through regression analysis that compared average classroom

scores among treatment teachers to those among control teachers, controlling for their students' characteristics and allowing for the clustering of teachers within schools. The study found no impacts on students' writing conventions, writing quality, or overall writing scores, although readers should interpret this finding with caution due to the small sample sizes and the low power of this analysis.

D. Potential explanations for findings

The patterns of findings over the first and second years of the study show that, while impacts on teacher practices appeared during the first year of engagement with TPP, impacts on students did not appear until after the first year. There are several possible explanations for the lag in impacts on student achievement. It is possible that teachers needed longer cumulative or sustained exposure to PD to impact students; teachers needed time to fully digest and implement what they learned; the changes in teachers' instructional practices during the first year were not large enough to affect student achievement that year; teachers did not experience impacts on their instructional practices early enough in the school year to affect student achievement; or teacher survey and observation instruments did not capture some aspects of practice that are important for student achievement.

Another notable pattern of findings was the lack of impacts on novice teachers' students, despite impacts on these teachers' instructional practices. This pattern is consistent with an explanation that it takes even more time for professional learning supports to affect novice teachers' students than to affect more experienced teachers' students.

These potential explanations for some of the study findings are no more than hypotheses. The study was not designed to identify the mechanisms by which TPP achieved impacts on teachers and students. Future work to investigate these hypotheses could advance our understanding further.

This page has been left blank for double-sided copying.

I. INTRODUCTION

Schools and teachers aim to prepare students to succeed in the classroom and ultimately in college and the workforce. Rigorous state standards, such as the Common Core State Standards (CCSS), are intended to further this goal by defining the core knowledge and skills students need to become college- and career-ready. As a vast majority of states have implemented CCSS, there is a growing need to help students develop these skills, and this hinges on ensuring that teachers are adequately prepared to provide the effective instruction in these areas. This need is even greater for novice teachers, who face the dual challenge of becoming effective teachers and meeting the new standards. It is important that all teachers, and especially novices, have effective supports that can build their capacity to use effective instructional practices to deliver high-quality material.

Integrating teacher professional learning with educative curriculum has emerged as a promising way to both build teacher capacity and support student achievement. In particular, this type of approach helps make professional development supports more relevant to teachers' specific contexts and thus easier to use, which could potentially address the finding that much of the current professional development does not meet teachers' needs (Bill & Melinda Gates Foundation 2014; Calvert 2016) and may not be associated with improvement in teacher performance (TNTP 2015).

EL Education designed the Teacher Potential Project (TPP) to build the instructional capacity of English language arts (ELA) teachers, and novice ELA teachers in particular. TPP provides a combination of a CCSS-aligned ELA curriculum and an embedded model of extensive professional development. TPP aims to enhance ELA teachers' instructional practices in areas aligned to CCSS and to foster student learning and achievement outcomes.

In 2013, EL Education was awarded a five-year, $11.9 million Investing in Innovation validation grant by the U.S. Department of Education. The purpose of the grant was to help build upon the evidence for EL Education's curriculum and professional development model by assessing the effectiveness of TPP. Mathematica assessed the implementation of TPP and its impacts on teacher practices and student achievement as the third-party evaluator for the grant. The study of TPP makes several important contributions to the literature evaluating paired curriculum and PD programs: It uses rigorous group designs, evaluates the impact of one and two years of program implementation, and examines broad outcomes on both teacher instructional practice and student ELA achievement. This report describes Mathematica's Evaluation of the Teacher Potential Project and its findings.

A. The Teacher Potential Project

EL Education's approach is based on the premise that, in order for schools to transform into dynamic communities of learning, all members of the school staff and student body—students, teachers, and school leaders—need to be actively engaged and committed to a cycle of ongoing learning. Building on this approach, TPP supports teachers to deliver high-quality ELA

instruction through two components: an ELA curriculum paired with teacher professional development (PD) learning supports.

- **ELA curriculum**. TPP provides an open-source ELA curriculum for grades 3 through 8, available online, that is aligned to CCSS. The curriculum is supported through classroom instruction where teachers serve as guides to student learning and emphasize discovery, critical thinking, and student collaboration.

- **Teacher professional development learning supports.** To support teachers' implementation of the curriculum, TPP provides a range of PD activities for ELA teachers and school leaders. These activities include on-site institutes where teachers can deepen their content knowledge and expertise. They also include on-site and video-based coaching from EL Education coaches that includes observations, modeling, lesson studies, and fostering professional learning communities within the school. Finally, they include access to online supports, including a range of online PD materials and opportunities to participate in online communities of practice and interactive webinars.

These components, and how they were implemented in this study, will be described in more detail in Chapter IV.

The goal of TPP is that novice ELA teachers (defined in this study as those in those with zero to three years of full-time teaching experience)—and all ELA teachers, regardless of their level of experience—will implement the curriculum and the instructional practices through their active engagement with TPP, leading to teaching practices aligned to CCSS and increases in students' ELA and writing achievement (see Appendix A).

B. Study design overview and research questions

Mathematica's evaluation of TPP consists of two study components: an implementation evaluation and an impact evaluation.

The **implementation evaluation** supports replication and scaling of TPP. Its findings will also help interpret the impact study findings by indicating the extent to which participants took up and engaged in the TPP services. The implementation evaluation consists of an analysis that identifies the core components of the program, defines the necessary implementation thresholds for these core components, and assesses the extent to which TPP was successfully implemented in treatment schools. These analyses use data collected through coach and teacher reports and classroom observations. It is guided by the overarching research question:

- Is there evidence of intervention fidelity?

The **impact evaluation** validates and expands the evidence base for EL Education's model of curriculum and PD. The impact evaluation was designed as a within-district school-level randomized controlled trial (RCT) to estimate the impacts of one year of TPP implementation on

teacher instructional practices and student ELA achievement. The RCT was conducted in 70 schools in 18 districts across 12 states from the 2014–2015 school year through the 2017–2018 school year. These findings are supplemented by a quasi-experimental design (QED) study to assess impacts of two years of implementation of TPP among selected schools on these teacher and student outcomes. Teacher outcomes are measured through surveys and classroom observations, while student outcomes are measured by standardized scores on state ELA assessments for all study students in both years as well as scores on a literacy task designed and implemented for this study for a subset of study students in spring 2017.

The impact evaluation is guided by two overarching research questions:

- What is the impact of TPP on the ELA instructional practices of novice teachers (defined in this study as those in those with zero to three years of full-time teaching experience) and all teachers in upper elementary and middle school grades?

- What is the impact of TPP on the ELA achievement of upper elementary and middle school students in the classrooms of novice and all teachers?

Within these general research questions, we investigate a number of specific questions, as shown in Table I.1 and described below.

In the evaluation of impacts of TPP on teacher practices, we examine impacts for teachers who experienced one or two years of TPP supports to understand how these different durations of engagement might change practices. We examine these impacts separately for novice ELA teachers as well as all ELA teachers to learn how TPP supports affect not just teachers in general, but specifically teachers who are relatively new to the profession. The impacts address research questions T1 through T4.

In the evaluation of impacts of TPP on student achievement, we examine a variety of impacts to understand how exposure to different types of engagement affects students. All impacts but one are on standardized ELA test scores. First, we examine impacts on students in the classrooms of study teachers during their first year of engagement with TPP, separately for the students of novice ELA teachers and all ELA teachers, to learn how one year of the curriculum and teacher supports might impact student ELA learning by the end of that year (research questions S1 and S2). To assess such impacts on student writing, an important facet of TPP's ELA curriculum, we also examine the performance of students in novice teachers' classrooms on a literacy task (research question S3). Next, we turn our attention to impacts that may be observed over a longer term. We examine impacts on students of study teachers in all study schools at the end of the schools' second year in the study—after some schools have engaged in just one year of TPP implementation and other schools have engaged in two years of implementation—to learn about the overall effects on students of at least one year of engagement with TPP (research question S4). To understand more fully the effects of two years of engagement with TPP, we examine impacts on students in the classrooms of teachers (novices and all teachers separately) who are in their second year of engagement with TPP (research questions S5 and S6). Because TPP includes

Table I.1. Impact evaluation research questions

	Research questions about impacts on teachers	Teacher impact estimate answering research question	Chapter in which impacts are reported
	What is the impact of the Teacher Potential Project on ELA instructional practices of 4th- through 8th-grade …		
T1.	novice ELA teachers who experienced the program for one year, compared to novice teachers in the control condition?	One-year novice impact	Ch. V
T2.	ELA teachers who experienced the program for one year, compared to teachers in the control condition?	One-year impact	Ch. V
T3.	novice ELA teachers who experienced the program for two years, compared to novice teachers in the control condition?	Two-year novice impact	Ch. V
T4.	ELA teachers who experienced the program for two years, compared to teachers in the control condition?	Two-year impact	Ch. V
	Research questions about impacts on students	**Student impact estimate answering research question**	**Chapter in which impacts are reported**
	Impacts after one year of school participation in the study		
	What is the impact of the Teacher Potential Project on ELA achievement on state assessments of 4th- through 8th-grade students who experienced the program for one year …		
S1.	with a novice ELA teacher, compared to students with a novice ELA teacher in the control condition?	One-year novice impact	Ch. VI.A
S2.	with an ELA teacher, compared to students in the control condition?	One-year impact	Ch. VI.A
S3.	What is the impact of the Teacher Potential Project on writing achievement (as measured by a reading/writing task) of 4th- through 8th-grade students of novice ELA teachers who experienced the program for one year, compared to students with a novice ELA teacher in the condition?	One-year literacy impact	Ch. VI.B
	Impacts after two years of school participation in the study		
	What is the impact of the Teacher Potential Project on ELA achievement on state assessments of 4th- through 8th-grade students who experienced the program for one year …		
S4.	in the year following their school's first year of program implementation, compared to students in the control condition?	Overall two-year impact	Ch. VI.A
S5.	with a novice ELA teacher who experienced the program for two years, compared to students with a novice ELA teacher in the control condition?	Two-year novice impact	Ch. VI.A
S6.	with an ELA teacher who experienced the program for two years, compared to students in the control condition?	Two-year impact	Ch. VI.A
S7.	during their school's second year of program implementation, compared to students in the control condition?	Two-year schoolwide impact	Ch. VI.A
S8.	with a novice ELA teacher in the year following their school's first and only year of program implementation, compared to students in the control condition?	One-year follow-up novice impact	Ch. VI.A
S9.	with an ELA teacher in the year following their school's first and only year of program implementation, compared to students in the control condition?	One-year follow-up impact	Ch. VI.A

ELA = English language arts.

supports aimed at school leadership as well as teachers, we also examine impacts on students in all ELA classrooms during the school's second year of engagement with TPP, without restricting our focus to teachers who were directly involved with the program (research question S7). Finally, we investigate whether the effects of TPP linger after teachers (novices and all teachers separately) stop receiving TPP supports by examining impacts on students in their classrooms in the year after they engaged with TPP (research questions S8 and S9).

C. Prior research on EL Education and TPP

TPP's combination of ELA curriculum and PD supports uses literacy practices, strategies, and protocols that are based on EL Education's approach to student literacy with network partners who adopt their whole-school model. The small but growing body of research on EL Education's whole-school model has shown its potential effects on student achievement. Using a quasi-experimental study design, a prior Mathematica study found that, compared with local district schools, the EL Education model implemented in five middle schools had a statistically significant positive impact on reading test scores after one year and statistically significant positive impacts on reading and math test scores after two and three years (Nichols-Barrer and Haimson 2013).

Other prior research has also shown promising results for the effectiveness of EL Education's model on student achievement, although these findings were based on studies with less rigorous research designs. Two studies found positive impacts on reading test scores (with one also finding impacts on math test scores) of students in EL Education schools, but they did not show that EL Education students were similar to comparison group students at baseline, so those impacts cannot be attributed solely to EL Education (Amoruso et al. 2010; UMass Donahue Institute 2011). A meta-analysis found that nine descriptive studies of student achievement at EL Education schools reported potential positive impacts of EL Education schools relative to non-EL Education schools averaging 0.19 standard deviations across all subjects and outcomes studied (Borman et al. 2001). However, these studies were all based on weaker research designs that inhibit the ability to make causal claims about evidence of effectiveness. For example, some studies did not control for baseline differences between groups, while others did not use a comparison group design.

Through the evaluation of TPP, this study is the first to assess the effects of an intervention based on EL Education's model using a randomized controlled trial design, the gold standard for evaluating program impacts. It is also the first to explore the impact that an intervention based on this model has on teachers' instructional outcomes. Earlier findings from this evaluation indicated that teachers who engaged with TPP demonstrated CCSS-aligned instructional practices to a greater extent than control teachers. Moreover, TPP impacted both novice and experienced teachers. At the end of one year of TPP, novice TPP teachers experienced impacts on specific practices related to developing student close reading and content knowledge, encouraging higher-order thinking skills, and asking students to cite evidence from texts (Choi et al. 2017). Teachers who engaged with TPP for two years also experienced impacts in encouraging higher-order thinking skills and asking students to cite evidence from texts; in

addition, their students engaged significantly more often in reading, writing, and speaking about texts in the second year of TPP (Choi et al. 2018).

D. Roadmap to the report

The remainder of this report describes the Evaluation of the Teacher Potential Project and its findings. Chapter II presents the study design and methods. Chapter III describes data collection methods and instruments. Chapter IV describes the analysis and findings of the implementation study. Chapter V presents the analysis and findings of the teacher impact evaluation, and Chapter VI presents the analysis and findings of the student impact evaluation. Chapter VII synthesizes and discusses the study findings about implementation and impacts. Appendices provide supplementary information for each chapter.

II. STUDY DESIGN AND METHODS

The Evaluation of the Teacher Potential Project includes two separate designs: a randomized controlled trial and a two-year quasi-experimental design. This chapter describes these designs and presents an overview of the approaches to evaluating implementation and impacts.

A. RCT study design

The impact evaluation used an RCT design to assess the impacts of a single year of TPP implementation on the outcomes of teachers and students in the study during that year. We also use this design to assess impacts on student achievement among all schools in the study—both those that had one year of TPP implementation and others that had two years of implementation—after two years, as described in Chapter VI.

1. Recruitment

The study team recruited school districts in urban, rural, and suburban locations across the United States. We focused on relatively high-need districts where at least 40 percent of students were eligible to receive free and reduced-price lunch. Districts were recruited in three cohorts during the school year prior to the intervention school year in which the district would participate. Cohort 1, which served as a pilot for the study, participated during the 2014–2015 school year. Cohorts 2 and 3 participated during the 2015–2016 and 2016–2017 school years, respectively. (Some schools in Cohort 3 participated during the 2017–2018 school year as well, as described in Section B.)

Within each cohort, participating districts selected elementary and middle schools meeting the following eligibility criteria to participate in the study:

- Schools have at least one novice teacher (defined in this study as those in those with zero to three years of full-time teaching experience) who teaches ELA.
- Schools serve grades between 3 and 8.
- Schools have at least 45 minutes of ELA classes.
- Schools were not currently using EL Education's grades 3 through 8 ELA curriculum as a primary instructional resource.
- School leaders and ELA teachers would be able to participate in the professional learning supports (for example, the school could secure substitute teachers if necessary and did not have union restrictions that prevent teachers from attending institutes in the summer).

A school's participation in the study involved being randomly assigned to either receive TPP at low or no cost for one year or continue with its existing ELA curriculum and teacher supports for

one year, as well as participation in student and teacher data collection for one year.[2] Districts in Cohorts 2 and 3 were later invited to provide student data for the subsequent school year. Cohort 3 districts were further invited to have teachers take surveys and have their classrooms observed in the fall and spring of the subsequent school year.[3] After contacting over 300 school districts, 19 districts initially agreed to participate and identified 79 schools to participate in the study. Table II.1 shows the numbers of school districts participating in each cohort.

Table II.1 School cohorts recruited for the study

	School years			
	2014–2015	2015–2016	2016–2017	2017–2018
Cohort 1 pilot	1 district			
Cohort 2		8 districts[a]		
Cohort 3, schools receiving one year of TPP			6 districts[b]	
Cohort 3, schools receiving two years of TPP			5 districts	5 districts

Note: Only participating districts are included; one district that participated in random assignment and subsequently dropped out is not included. Shaded cells show the school year in which treatment schools in each cohort received TPP services.

[a]Includes the Cohort 1 pilot district.

[b]Includes one Cohort 2 district.

2. Random assignment

Schools were randomly assigned to treatment and control conditions within each school district as the districts were recruited.[4] In two cases, random assignment was conducted in separate rounds in a district, as they identified additional schools to participate. Schools were selected as the unit of assignment because TPP includes some whole-school elements, making random assignment of teachers or classrooms within schools infeasible. Given the relatively small number of study schools within each district, random assignment was conducted within matched school pairs to help ensure balance across the treatment and control groups. Pairs were constructed for elementary schools and middle schools separately. Matching was based on data on school characteristics, including grade configuration, numbers of teachers and students, and student demographics and test scores, as well as information from principals on the number of expected ELA novice teacher hires in the upcoming school year. Within each pair, one school was assigned to treatment condition and one to control, so that each school had a 50 percent probability of assignment to each condition. In districts with an odd number of study schools, the unpaired school was assigned to either treatment or control condition with a 50 percent

[2] Two districts contributed a small amount to receive TPP services; the remaining districts received services at no cost.

[3] Cohort 2 districts were not invited for extended teacher data collection.

[4] If a state had two participating districts that each identified one school for inclusion in the study, we randomly assigned one district to the treatment condition and one to the control condition. This occurred in two states—one in Cohort 2 and one in Cohort 3.

probability of assignment to each. Appendix B provides additional details about the random assignment procedure.

Of the 79 schools in 19 districts, 40 were assigned to the treatment condition and 39 were assigned to the control condition. Table II.2 shows the numbers of schools randomly assigned overall and within each cohort. The one Cohort 1 pilot district contained two schools randomly assigned to the treatment and control conditions. Twenty-two schools in 8 districts were randomly assigned in Cohort 2; the 2 schools in the Cohort 1 pilot district continued to participate in another year of the study after the pilot year, bringing the total of Cohort 2 schools randomly assigned to 24. Fifty-five schools in 11 districts were randomly assigned to Cohort 3, including additional schools in one district that contributed schools to Cohort 2.

Table II.2. Numbers of schools in the randomized controlled trial study

	Number of schools randomly assigned	Number of schools participating in the randomized controlled trial study
All schools		
Treatment	40	35
Control	39	35
Total	79	70
By cohort		
Cohort 1 pilot (started 2014–2015)		
Treatment	1	1
Control	1	1
Total	2	2
Cohort 2 (started 2015–2016)[a]		
Treatment	12	9
Control	12	9
Total	24	18
Cohort 3 (started 2016–2017)		
Treatment	28	26
Control	27	26
Total	55	52

[a] Includes the two Cohort 1 pilot schools.

3. Implementation of experimental conditions and teacher eligibility

After random assignment was conducted, the study team communicated school experimental assignments to districts. Schools were typically notified of their assignment between February and June of the school year prior to the intervention year (although the pilot schools in Cohort 1 were notified in August). This early notification was necessary so that EL Education could begin planning for the delivery of the intervention, which included summer activities, with the treatment schools. As part of this process, EL Education and the study team identified eligible teachers who would be considered study participants at each school, both treatment and control. We call these teachers "study teachers." Using rosters of teachers in the study schools collected

from districts, we identified eligible teachers as those teaching ELA classes that were not reading-only and did not exclusively serve special student populations, such as ELL or special education. We further used the rosters to identify which eligible teachers were novices. Teachers who joined schools after the rosters were received in the fall did not participate in TPP or the study.

4. Description of participating schools

Some districts and schools dropped out after random assignment, leaving 70 participating schools in 18 districts, as shown in Table II.2.[5] One school district in Cohort 2 with 6 study schools (3 treatment and 3 control) dropped out of the study after receiving their school assignments. Three other schools in Cohort 3, a control school from one district and a treatment and control school (not a matched pair) from another district, also dropped out of the study after random assignment.[6] The treatment school had already made plans for teacher PD in the upcoming school year and dropped out before receiving any TPP services; the two control schools were not able to fulfill data requests, one due to demands from the district related to a separate initiative. The remaining participating schools and teachers complied with their experimental assignments. There was no evidence that treatment schools did not implement TPP (see Chapter IV) or that control schools implemented elements of TPP. Furthermore, there is no evidence of teacher crossover, either due to teachers moving to a school with a different experimental assignment or due to teachers not complying with their experimental assignment (for example, treatment teachers not implementing at least some elements of TPP or control teachers implementing elements of TPP).

The participating schools had a relatively low socioeconomic status, serve a largely minority race and ethnicity population of students, and have below-average student achievement. Among the participating schools, 71 percent of students were eligible for free or reduced-price lunch, 27 percent were white non-Hispanic, 46 percent were black non-Hispanic, 21 percent were Hispanic, 6 percent reported another race, and 48 percent were female (see Table II.3). The schools had approximately 531 students enrolled and 36 teachers on average. Average student standardized ELA test scores from the year prior to the intervention were –0.5.[7] There were no statistically significant differences in these features between treatment and control schools. This suggests that random assignment achieved balance across the experimental groups, at least in terms of observed characteristics.

[5] Several schools split into two after agreeing to participate: two Cohort 2 control schools, one Cohort 2 treatment school, and two Cohort 3 treatment schools. Each of these schools is counted as a single school in this chapter but will be counted separately in the discussion of teacher and student impacts in Chapters V and VI.

[6] The schools matched to these three schools were not dropped from analysis.

[7] The average student standardized score among the norming population is zero.

Table II.3. Characteristics of schools participating in the randomized controlled trial study

	Treatment mean (SD)	Control mean (SD)	p-value of difference	Overall mean (SD)
Whether charter school (percent)	5.3 (0.23)	0.0 (0.00)	0.17	2.7 (0.16)
Percent students eligible for free or reduced-price lunch	71.5 (0.26)	69.5 (0.22)	0.72	70.6 (0.24)
Student race and ethnicity				
Percent white non-Hispanic	22.4 (0.28)	32.2 (0.33)	0.17	27.2 (0.31)
Percent black non-Hispanic	49.0 (0.34)	42.5 (0.33)	0.40	45.8 (0.34)
Percent Hispanic	21.7 (0.23)	19.5 (0.19)	0.65	20.7 (0.21)
Percent other race	6.9 (0.08)	5.8 (0.05)	0.47	6.3 (0.07)
Percent students female	48.3 (0.03)	48.4 (0.02)	0.94	48.3 (0.03)
Number of students enrolled	521.1 (214.48)	541.5 (202.14)	0.68	531.0 (207.39)
Number of full-time teachers	35.9 (15.18)	35.8 (11.39)	0.98	35.8 (13.38)
Average student standardized ELA pre-test scores	–0.51 (0.47)	–0.48 (0.43)	0.71	–0.50 (0.45)
School sample size	38	36		74

Source: Common Core of Data (CCD), years 2015–2016, except ELA pre-test scores standardized by national-level means and standard deviations (see Chapter VI), which come from the sample of student administrative data collected from schools. For seven schools, this year of data was not available and year 2016–2017 was used. One school did not have any CCD data available. For four schools, school free or reduced-price lunch data were collected from the district. The table includes 74 schools rather than 70 because several schools split in two after random assignment and CCD data were only available for the separate schools. Among the control group, two schools split in two and another had no CCD data available, so the number of control schools is 36 instead of 35. Among the treatment group, three schools split in two, so the number of treatment schools is 38 instead of 35. The control school that had no CCD data available did contribute ELA pre-test score data, so for that measure 38 treatment and 37 control schools are included for a total sample size of 75.

ELA = English language arts; SD = standard deviation.

B. Two-year QED study design

In addition to the RCT, we designed a QED study to assess the impact of extending implementation of TPP to a second year. The motivation was that EL Education recommends teachers utilize the EL Education curriculum and professional learning supports for at least two years to see the full benefits of the intervention. Indeed, research suggests that it can take two or more years of teacher PD to see impacts on student outcomes (Glazerman et al. 2010; Schmidt et al. 2017).

During the winter of 2016, EL Education contacted each district in Cohort 3 to participate in a second year of the study in the following 2017–2018 school year (see Table II.1). Treatment schools that chose to participate for a second year would continue to implement TPP for a second year, while control schools continued to provide typical teacher supports and agreed not to use the TPP curriculum as a primary curriculum; both treatment and control schools would continue to participate in data collection for the study. Some districts declined to participate for a second year because schools would not be able to commit due to time constraints or shifts in district priorities. For example, one school was implementing an intensive school improvement program

and did not want to overburden their teachers and staff with multiple intensive initiatives. Some districts agreed to participate, but particular schools within those districts did not.

Twenty-two schools (10 treatment, 12 control; and 18 of which were matched pairs) in five districts agreed to participate in this second year, which we call the "two-year QED," as shown in Table II.4.[8] The two-year QED study has a quasi-experimental design because schools were purposively selected to participate based on the decisions of EL Education, districts, and schools. The study team collected rosters of teachers in study schools from districts at the beginning of the second year. Teachers who joined schools after the rosters were received in the fall did not participate in TPP. Eligible study teachers who stayed in study schools into the second year continued in the study. The components of TPP as implemented in the second year were generally similar to those implemented in the first year with some minor changes, as Chapter IV describes.

Table II.4. Numbers of participating schools by years of intervention of TPP

	Number of schools participating in the study
Schools in the one-year only study	
Treatment	25
Control	23
Total	48
Schools in the two-year QED study	
Treatment	10
Control	12
Total	22
Total schools in the RCT study	70

QED = quasi-experimental design; RCT = randomized controlled trial.

We sometimes characterize the remaining 48 schools from the RCT study as being in the "one-year only study," since the treatment schools in that group received just a single year of implementation of TPP.

The schools participating in the two-year QED study were somewhat lower in their socioeconomic status and had slightly higher baseline ELA achievement than the full sample of schools participating in the RCT. Among these schools, 61 percent of students were eligible for free or reduced-price lunch, 32 percent were white non-Hispanic, 40 percent were black non-Hispanic, 22 percent were Hispanic, 6 percent reported another race, and 49 percent were female (see Table II.5). The schools had approximately 556 students enrolled and 35 teachers on average. Average student standardized ELA test scores from the year prior to the intervention were –0.25. There were no statistically significant differences in these features between treatment and control schools.

[8] One of the treatment schools that split into two after agreeing to participate is in this sample of schools. This school is counted as a single school in this chapter but will be counted separately in the discussion of teacher and student impacts in Chapters V and VI.

Table II.5. Characteristics of schools participating in the two-year quasi-experimental design study

	Treatment mean (SD)	Control mean (SD)	*p*-value of difference	Overall mean (SD)
Whether charter school (percent)	0.0 (0.00)	0.0 (0.00)	.	0.0 (0.00)
Percent students eligibility for free or reduced-price lunch	59.5 (0.34)	61.6 (0.24)	0.86	60.6 (0.29)
Student race and ethnicity				
Percent white non-Hispanic	33.9 (0.29)	29.9 (0.27)	0.73	31.9 (0.27)
Percent black non-Hispanic	32.3 (0.27)	47.4 (0.30)	0.21	39.9 (0.29)
Percent Hispanic	28.0 (0.28)	16.5 (0.15)	0.22	22.3 (0.23)
Percent other race	5.9 (0.05)	6.2 (0.03)	0.85	6.0 (0.04)
Percent students female	49.4 (0.05)	48.9 (0.03)	0.73	49.2 (0.04)
Number of students enrolled	541.3 (186.80)	573.8 (177.82)	0.67	557.5 (179.13)
Number of full-time teachers	33.4 (8.70)	35.7 (9.73)	0.55	34.6 (9.10)
Average student standardized ELA pre-test scores	–0.24 (0.40)	–0.27 (0.39)	0.87	–0.25 (0.39)
School sample size	12	12		24

Source: Common Core of Data (CCD), years 2015–2016, except ELA pre-test scores standardized by national-level means and standard deviations (see Chapter VI), which come from the sample of student administrative data collected from schools. For two schools, this year of data was not available and year 2016–2017 was used. The table includes 12 treatment schools instead of 10 because two schools split into two campuses by grade level after random assignment and CCD data were only available for the separate campuses.

ELA = English language arts; SD = standard deviation.

C. Evaluation methods

To assess the fidelity of implementation of TPP, we developed measures of implementation fidelity for each TPP component based on whether schools delivered the ELA curriculum and whether and to what extent teachers participated in PD activities. We examined fidelity among novice and all teachers during the first and second year of implementation. Details of the specifications of each measure, as well as results, are described in Chapter IV.

We used a single general approach to estimate impacts of TPP on teacher and student outcomes at different time points to address each impact study research question, including for the RCT study and the two-year QED study. We estimated impacts by comparing average outcomes across the treatment and control groups, adjusting for characteristics of the sample that may be related to the outcomes and that may differ across experimental groups and allowing for the clustering of the sample within schools. The set of characteristics and other features of the models, such as the treatment of missing data and weights, differ across the analyses. Details of the specifications used in the analyses of impacts on teachers, as well as results, are described in Chapter V. Details and results of analyses of impacts on students are described in Chapter VI.

This page has been left blank for double-sided copying.

III. INSTRUMENT DEVELOPMENT AND DATA COLLECTION

Data for the Evaluation of the Teacher Potential Project was collected from several sources. This chapter presents a brief description of each data source and collection procedures.

TPP data collection comprised five components: (1) teacher rosters, (2) surveys (teacher surveys, professional development exit surveys, and TPP coach surveys), (3) teacher observations, (4) the student literacy task, and (5) student administrative records. Table III.1 provides an overview of the data collection components, which we discuss below in more detail.

Table III.1. Overview of data collection

Instrument	Respondents/ Sources	Data collection timing	Mode	Purpose	Study Component
A. Teacher rosters	Administrative data from schools or districts (treatment and control)	Fall 2015–2016; Fall 2016–2017; Fall 2017–2018	Electronic district files	Build the sample for the evaluation	Implementation evaluation, teacher impact evaluation, student impact evaluation
B. Surveys					
Teacher survey	Teachers (treatment and control)	Fall and spring 2015–2016; Fall and spring 2016–2017; Fall and spring 2017–2018	Web-based self-administered survey	Measure teachers' ELA instructional practices and document fidelity of TPP implementation	Implementation evaluation, teacher impact evaluation
Professional development exit survey	Teachers (treatment)	Summer and fall 2015–2016; Summer and fall 2016–2017	Web-based self-administered survey	Document fidelity of TPP implementation	Implementation evaluation
TPP coach survey and Teacher TPP participation form	TPP coaches	Fall and spring 2015–2016; Fall and spring 2016–2017; Fall 2017	Web-based self-administered survey	Document fidelity of TPP implementation	Implementation evaluation
C. Teacher observations	Teachers (treatment and control)	Fall and spring 2015–2016; Fall and spring 2016–2017; Fall and spring 2017–2018	On-site standardized observation	Measure teachers' ELA instructional practices and document fidelity of TPP implementation	Implementation evaluation, teacher impact evaluation
D. Student literacy task	Students of novice teachers in Cohort 3 schools (treatment and control)	Spring 2017	Teacher administered writing project	Measure student literacy achievement	Student impact evaluation
E. Student administrative records	Administrative data from schools or districts (treatment and control)	2014–2015; 2015–2016; 2016–2017; 2017–2018	Electronic district files	Measure student ELA achievement	Student impact evaluation

ELA = English language arts.

A. Teacher rosters

To identify the population of teachers for this study, the study team contacted the school districts for rosters with teacher administrative data for all the ELA teachers in grades 4 through 8 in the 70 participating schools. Eligible teachers included those who were full-time and teaching ELA classes that were not reading-only and did not exclusively serve special student populations, such as ELL or special education. The rosters, collected between October and December of each school's first year of participation in the study, were used to build the sample for the evaluation and included the following information: (1) teacher identification number, (2) teacher name, (3) school name, (4) grade level the teacher taught, and (5) teacher-preferred email address. We worked with districts to correct any missing data or discrepancies. We then reached out to an administrator at each school to request the number of years of teaching experience for each identified teacher in order to determine each teachers' novice or experienced status.

During the first two years of data collection (2015–2016 and 2016–2017 school years), the treatment teacher rosters and teachers' novice status were verified by providing the information to TPP coaches to review. For control schools, we reached out to a school contact at each school and asked them to verify teacher rosters and teachers' novice status. As a result of this process, we identified a final sample of 361 ELA teachers (199 control and 162 treatment) who were eligible to be part of the study, 113 of which were novice teachers (65 control and 48 treatment). These teachers comprise our sample of those who received one year of TPP.

It is important to note that because teacher roster verification occurred after random assignment occurred and schools were included in the study, this resulted in the study learning that some schools did not have teachers that met the study's definition of novice teachers (those with zero to three years of full-time teaching experience). This occurred because districts misidentified teachers who were new to the district as being new to the teaching profession. But our roster verification process identified that these particular teachers had full-time teaching experience in other school districts. TPP implementation and data collection still occurred in these schools as planned. As a result, there were 60 participating schools (31 control and 29 treatment) that had at least one novice ELA teacher for this study.

During the 2017–2018 school year, we verified the teacher rosters with TPP coaches (for the Cohort 3 schools receiving a second year of TPP) and with the school contacts at the corresponding control schools. In particular, we asked these individuals to verify which teachers from the 2016–2017 school year were still at the school and teaching ELA in grades 4 through 8. The confirmed teachers were then considered our sample of two-year TPP teachers (in other words, those who were direct recipients of two years of TPP). This included 55 ELA teachers (21 control and 34 treatment), of whom 23 were novice teachers (11 control and 12 treatment). For evaluation purposes, teachers classified as novice during their first year of the study maintained their novice status during their second year of the study. Of the 22 schools participating in the second year of TPP (12 control schools and 10 treatment schools), 12 schools still had novice teachers eligible for the study (5 control schools and 7 treatment schools).

B. Surveys

Three survey instruments were developed for the evaluation to collect information from those directly involved in the study: teachers and TPP coaches. These instruments include: (1) a teacher survey, (2) a professional development exit survey, and (3) a TPP coach survey. In this section, we describe the development of and data collection with these survey instruments.

1. Teacher survey

The teacher survey was primarily designed to measure teachers' self-reports of the types of instructional practices used and their PD experiences, to inform the implementation and teacher impact evaluations. The developed survey included items that focused on ELA instructional practices related to key CCSS shifts and general teacher instructional practices. There were also items that collected information on the general PD supports that teachers received and, for treatment teachers, those experienced as part of their participation in TPP. Table III.2 provides an overview of the teacher survey's main areas and the different topics covered within those areas (see Appendix E for more details on the methodology used to develop the teacher instructional practice analytic measures).

Table III.2. Teacher survey content

Main areas	Topics
General instruction	Content knowledge
	Higher-order thinking
	Multimedia use
	Prior knowledge
	Student independence
	Student responsibility for learning
Reading and writing instruction	Multiple types and purposes of writing
	Reading, writing, and/or speaking about texts
	Use of text evidence
	Writing conventions
Classroom management and environment	Self-reported classroom management
Professional development experiences	General professional development experiences
	TPP professional development experiences

After the teacher survey was pre-tested with the Cohort 1 pilot schools in the fall and spring of the 2014–2015 school year, the teacher survey was administered to all eligible teachers two times during the school year: in the fall and spring of the 2015–2016 (for Cohort 2 schools), 2016–2017 (for Cohort 3 schools), and 2017–2018 school years (for Cohort 3 schools in their second year of TPP implementation). The field period lasted for about three months for each data collection period (fall data collection occurred from November through January and the spring

data collection occurred from March through May). Teachers received a $50 electronic gift card to thank them for completing the survey. Appendix C presents a copy of the final teacher survey.

2. Professional development exit survey

The professional development exit survey measured the participants' perceptions of the usefulness of the TPP PD learning institutes delivered by EL Education staff and coaches. Teachers and school leaders were asked to complete the online 10-minute survey at the end of each learning institute they attended during the summer and fall of the 2015–2016 and 2016–2017 school years. The survey contained items that asked respondents to rate the usefulness and delivery of each institute and open-ended questions on what the respondents learned from the institute and how they would apply that information. In addition to supporting the study, the survey provided TPP coaches and EL Education with information on how well the institutes were being implemented to help improve these PD supports.

3. TPP coach survey and teacher TPP participation form

All treatment schools in the study were provided an EL Education coach who provided PD on-site and virtually, and supported the implementation of the TPP ELA curriculum and instruction at the schools. The TPP coach survey asked coaches about which PD supports they provided and their perceptions of successes and challenges related to implementation. The coaches completed a survey for each of the schools to which they were assigned. In addition to supporting the study, the survey provided EL Education with information to help improve PD services and implementation, as well as gauge the level of implementation of TPP in the treatment schools. Coaches completed the TPP coach survey at the end of each semester in the fall and spring of the 2015–2016 (for Cohort 2 schools) and 2016–2017 (for Cohort 3 schools) school years, and the fall of the 2017–2018 school year for Cohort 3 schools in their second year of implementation of TPP. Coaches also completed a teacher-specific form for each of the teachers for whom they provided PD supports. This form was designed to capture each teacher's level of engagement with specific supports, such as their participation in learning institutes and the number of meetings they had with their TPP coach.

C. Teacher observations

The study team designed an observation instrument specifically for this study that systematically captured information about treatment and control teachers' practices, for those aligned with CCSS in ELA. Similar to the teacher survey, the observation instrument included items related to teachers' general instruction, reading and writing instruction, and classroom management and environment. Table III.3 shows the observation instrument's main areas and the different topics covered within those areas (see Appendix E for more details on the methodology used to develop the analytic measures of teacher instructional practices). The observation instrument was reviewed by experts in ELA instruction and then piloted with Cohort 1 schools and other schools not in the study's sample during the fall and spring of the 2014–2015 school year. Appendix D presents a copy of the final teacher observation instrument.

Table III.3. Teacher observation instrument content

Main areas	Topics
General instruction	Connected lessons
	Connections to world
	Content knowledge
	Higher-order thinking
	Multimedia use
	Student participation
Reading and writing instruction	Academic vocabulary
	Multiple types and purposes of writing
	Reading, writing, and/or speaking about texts
	Use of text evidence
Classroom management and environment	Classroom climate
	Classroom management

Two members of the study team conducted an intensive training to train observers to reliably observe classroom instruction. All observers participated in an intensive training at the end of the summer prior to the start of each school year (2015–2016, 2016–2017, and 2017–2018). As part of the training, observers used the observation instrument on videos of classroom instruction as practice and then applied the instrument on videos of classroom instruction to simulate a classroom observation they would conduct when in the field. After the first round of observations was completed each school year, observers underwent a refresher training to go over each item of the observation instrument. Before and after each observation in the field, observers used the instrument on videos of classroom instruction to ensure that they were answering questions accurately. The percent of agreement between the observers on the observational instrument was regularly calculated throughout the data collection field periods, and the average interrater reliability across all data collection periods was 0.83.[9]

The classroom observations occurred in both fall and spring of particular school years with field periods that paralleled the teacher survey administration. Observations were conducted in the 2015–2016 school year for Cohort 2 and in the 2016–2017 and 2017–2018 school years for Cohort 3. Teachers' ELA instruction practices were observed in the classroom for a single class period, which tended to range from 45 to 60 minutes depending on the school schedule.

D. Literacy task

To collect data on the CCSS-aligned writing of novice teachers' students, the study team developed a student literacy task that consisted of an opinion writing (for grades 4 and 5) or an argument writing (for grades 6 to 8) activity. For grades 4 and 5 and 6 to 8, the literacy task

[9] Interrater reliability was 0.87 in fall 2015, 0.78 in spring 2016, 0.84 in fall 2016, 0.83 in spring 2017, 0.83 in fall 2017, and 0.84 in spring 2018.

prompted students to take a stance on an issue and write a letter to the principal to convince him or her of their stance using reasons and evidence from the text. The grade 4 and 5 task focused on whether the school should get a new lunch menu and included one informative text for students to read and use to provide reasons and evidence for their side. The grade 6 to 8 task focused on whether the school should adopt using video games in the classroom and included two informative texts for students to read and use to provide reasons and evidence for their side. In its development, the literacy task underwent expert review from four ELA teachers who reviewed and piloted measures in their grade 4–8 classrooms during the 2015–2016 school year; the scoring rubrics were tested in the summer of 2016 by three ELA teachers and two study team members. The scoring rubrics were revised and used to score the approximately 200 student writing samples from the pilot test.

Trained raters scored the student writing samples on a scale of 1 to 5. Student writing was scored on outcomes related to writing conventions, such as vocabulary and grammar, outcomes related to writing quality (such as evidence from texts and counterclaims), as well as overall literacy scores. Writing samples from students in grades 4 to 6 were scored on 15 outcomes and samples from students in grades 7 and 8 were scored on two additional outcomes: counterclaims and a total score that included counterclaims scores. Students in grades 7 and 8 were asked to use counterclaims in their writing because this standard is introduced for students by CCSS beginning in grade 7. Table III.4 provides a description of each outcome organized by the writing conventions, writing quality, and overall literacy areas. See Appendix G for more information about the literacy task measure.

Data in the form of student writing samples from the literacy task were collected from novice teachers in Cohort 3 during the spring semester of 2017. To minimize bias that might be introduced when teachers select which classrooms to submit writing for, teachers were asked to send us the writing of students in their first ELA class of the morning and their first ELA class of the afternoon. Teachers or a school staff support member were asked to first obtain parental consent through hard copy. Teachers then administered the writing activity in their classes and provided students 45 minutes to complete the activity in full. Each teacher who sent back 80 percent or more parental consent forms received a $50 gift card. Teachers were given a second gift card of $50 once they provided their students' writing data. The interrater reliability of the returned student writing that was scored had an exact or adjacent agreement overall of 0.81.

Table III.4. Description of literacy outcomes

Domain	Outcome	Description of outcome
Writing conventions	Vocabulary	Appropriate use and variety of vocabulary that are both from the provided text(s) and outside the texts.
	Mechanics	Accuracy of mechanics usage and clarity of writing related to mechanics usage.
	Spelling	Accuracy of spelling and clarity of the writing related to spelling.
	Grammar	Accuracy of grammar use and clarity of the writing related to grammar usage.
	Writing conventions total	Sum of vocabulary, spelling, grammar, and mechanics scores.
Writing quality	Voice	Writer's clarity, tone, and style. The extent to which the writer uses tone and clarity to assert their position clearly.
	Sentence variety	Variety of length and structure of sentences.
	Organization	Structure supporting the effectiveness of the writing. The extent to which the writing has a beginning paragraph with a thesis statement, at least one body paragraph, and an ending paragraph with a restatement of the thesis.
	Statement of position	The extent to which the writer stated a relevant position with reasons to support the position.
	Evidence	The extent to which the writer used relevant information from the text to support their reasons.
	Reasons	The extent to which the writer included reasons that were relevant to the stated position that were from the texts.
	Transition words and links	The extent of the student's used linking words (grades 4–5) and/or transition words, phrases, and clauses (grades 6–8) to create cohesion and clarify the relationship between the position, reasons, evidence, and/or counterclaims (grades 7–8) throughout the writing.
	Counterclaims	Student's statement of the opposing position, statement refuting the opposing position, and explanation of the logic using information from the text. Per CCSS, this was only requested of and scored for students in grades 7–8.
	Writing quality total	The sum of voice, sentence variety, organization, statement of position, reasons, evidence, and linking or transition words and links
	Writing quality total (grades 7 and 8)	The sum of voice, sentence variety, organization, statement of position, reasons, evidence, linking or transition words and links, and counterclaims.
Overall literacy	Total score	The sum of writing conventions total and writing quality total.
	Total score (grades 7 and 8)	The sum of writing conventions total and writing quality total (grades 7 and 8).

CCSS = Common Core State Standards.

E. Student administrative records

There were three types of student administrative data that we requested from each of the districts for all treatment and control schools in the study: (1) student roster data, (2) student background data, and (3) student test score data. We describe these types of data below. Across each of these types of data, we requested student and school identifiers for linking purposes. We asked district data administrators to provide prior year and current year data for the 2015–2016 (for Cohort 2 schools), 2016–2017 (for Cohort 3 schools), and 2017–2018 (for Cohort 3 schools in their

second year of TPP implementation) school years. All 18 participating study districts provided at least some of the requested data, but not all could be included in the student impact analysis.[10]

1. **Student roster data.** We collected roster data for all students in grades 4 through 8 (depending on the grade configuration of the school) for the schools' first year of participation in the study and the previous school year for the analysis of one-year student impacts. We also requested data for the schools' second year of participation in the study for the analysis of two-year student impacts. These data, when matched to study teacher rosters, allowed us to identify all students in the eligible classes of study teachers. The student roster data also specified the class periods or sessions in which students received ELA instruction. The specific roster data elements included a student identifier, teacher name and identifier, school name and identifier, district name, grade, and class name or identifier. We requested student roster data for the end of October and the end of the year for each school year of school participation. We allowed districts to provide rosters from any time between the start of the school year and the end of October if it was infeasible to provide a roster for the end of October.[11]

2. **Student background data.** For the students listed in each of the roster files, we requested data on their demographic and background characteristics. These data were needed to provide descriptive information of our sample of students and to control for these factors and their potential association with outcomes in the impact analysis. The specific data elements included students' grade, gender, race and ethnicity, English-language learner status, special education status, disability code, and free and reduced-price lunch eligibility status.

3. **Student test score data.** We requested test scores on the ELA and math tests administered by states for the students included in the roster files. We used the ELA test score for a specified school year of school participation as a student achievement outcome measure and the ELA test score from the previous school year as a pre-test measure. The specific data elements included test name, grade level, testing date, scaled score, retake code, and exemption code.

[10] Student administrative data from two districts (for four schools) was not included in our analysis. One district provided their data past the set deadline needed for its inclusion in this evaluation. Another district did not provide student administrative data for all the requested school years needed for the evaluation.

[11] The study team had originally planned to obtain the student rosters before providing schools' experimental assignments to districts. This plan would ensure that the assignment of classes or students to teachers would not be influenced by the school's assignment. For example, school leaders might be more willing to assign challenging classes or students to novice teachers if they knew the novices would be receiving the support of TPP. This plan proved to be infeasible since class rosters were typically not finalized until after the school year began, and schools had to receive their experimental assignments well before the start of the school year. Despite this timing, we believe it is unlikely that school assignments influenced the composition of class rosters. There are typically other high-priority objectives that rosters are designed to achieve, such as student balance, which may not leave much opportunity for this discretion.

IV. TPP PROGRAM IMPLEMENTATION AND FIDELITY ASSESSMENT

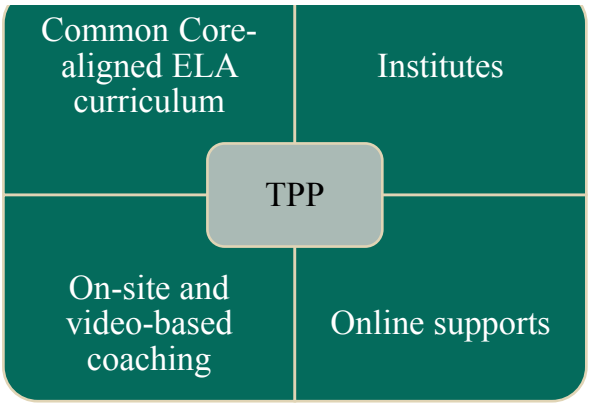
To provide context for understanding the study and ultimately interpreting the study's impact findings, a description of the TPP program and its implementation is essential. To support replication of TPP in the future, it is important to assess the fidelity of the implementation. In this chapter, we describe the TPP program and its implementation in this study and present the results on the extent to which TPP was implemented with fidelity. First, we discuss the components of the program and how they were delivered based on information gathered from program documents and discussion with EL Education leadership. Second, we present the findings from the implementation evaluation that assesses the fidelity of implementation of TPP. The implementation evaluation involved first developing measures of implementation fidelity for each TPP component based on quantifiable aspects of program delivery, then examining these among novice and all teachers during the first and second year of implementation as evidence of intervention fidelity. The measures were based on analysis of data from coach surveys, professional development exit surveys, classroom observation, and teacher surveys to determine whether TPP teachers engaged each of the program's key components to a high degree.

A. Description of the TPP program

TPP features a unique blend of a Common Core-aligned English language arts curriculum and PD that include institutes, personalized on-site and video-based coaching, and online supports. These structured and comprehensive resources and professional learning opportunities are designed to help prepare teachers to become more effective educators and align their instruction to the rigorous objectives of CCSS.

Below we describe these TPP components and what they included during schools' first year of implementation (see Figure IV.1), along with changes that were made to the components in the following year for schools receiving a second year of TPP.

Figure IV.1 Key components of the TPP program



1. Components

The **Common Core-aligned ELA curriculum**, designed for grades 3 through 8, is aligned to the instructional demands of CCSS that aim to prepare students for college and the workplace. The curriculum consists of Common Core-aligned literacy modules, which span eight to nine weeks of instruction and are designed to have students engage with complex texts and write informational, argumentative, and narrative texts intended for external audiences. There is also a strong emphasis on giving students the opportunity to work together and collaborate. Each

module consists of daily lesson plans, curriculum maps, learning targets and performance tasks, aligned assessments, and extensions for students needing more challenging work or remediation.

The **institutes** are interactive PD workshops with a TPP coach or other EL Education staff; they can last one to three days and are designed to deepen teachers' content knowledge and instructional expertise. These are role-specific institutes that help ensure that participants receive both critical, foundational learning about CCSS and its implications for classroom practice, as well as support with initial plans. There were four TPP institutes offered to teachers during the summer and the school year: a Meet the Modules Institute and three Instructional Institutes. A Leading Implementation Institute was also provided to school leaders and instructional coaches.

- The *Meet the Modules Institute*, a three-day introductory institute given during the summer, was open to ELA and specialist teachers who would use the EL Education curriculum. The purpose of this institute was to introduce teachers to the curriculum and its alignment to the instructional shifts of CCSS and how to implement the curriculum modules in their classrooms during the year.

- The three *Instructional Institutes* included "Using Data to Sharpen Implementation" (a two-day institute during the summer), "Supporting All Learners" (a two-day institute during the summer), and "Complex Texts in the Content Areas" (a one-day institute during the fall). These institutes were open to all teachers, including non-ELA teachers (such as science and social studies teachers, special education teachers, and teachers of English learners), and were designed to provide participants the opportunity to engage in pedagogical approaches needed for delivering the ELA curriculum.

- The *Leading Implementation Institute*, a two-day institute during the fall, was intended to provide school leaders and instructional coaches with tools to provide support and feedback to teachers as they implemented the curriculum. This institute helped attendees learn to coach to foster teachers' learning and how to shift their teaching practices to be aligned to the CCSS instructional shifts. The institute also provided school leaders with PD on to foster school-level changes that support teachers' professional growth through aligning teacher PD, feedback, and coaching with formal teacher evaluation systems.

On-site and video-based coaching was primarily provided to novice ELA teachers, although other ELA teachers also worked with TPP coaches. TPP coaches are professional development specialists who are extensively trained by EL Education to support school leaders and teachers in their implementation of the ELA curriculum and support teachers in their PD. The coaching provided to novice teachers consisted of observing teachers in their classroom, providing feedback to support their implementation of TPP's ELA curriculum, and modeling effective classroom instructional practices that are aligned to the curriculum and the CCSS instructional shifts. Coaches also supported novice teachers in unpacking the curriculum modules and engaged novice teachers in monthly lesson study sessions. In these lesson studies, coaches focused on best practices for delivering upcoming curriculum modules that aligned to CCSS

instructional shifts, such how to incorporate informational texts and have students engage in complex texts and read and write from evidence.

Coaches also helped foster professional learning communities (PLCs) within schools and met with school leadership. On a monthly basis, the PLCs would convene all teachers—novice and experienced and across different subject areas—and focus on student writing across the curriculum. After initiating and leading the first PLC, coaches focused on developing teacher leadership to sustain them and then observe and provide feedback on ensuing PLC meetings. During monthly school leadership planning sessions, coaches would work collaboratively with the school implementation teams to create an annual work plan and calendar to support novice teachers' PD, provide guidance on how to monitor progress toward implementation, and help ensure that schools' ELA- and CCSS-related objectives were being met. There were a total of 21 on-site coaching days, 3 video-based coaching days, and 1.5 days of strategic planning at each school.

Online supports provided to teachers included collaborative and independent professional learning activities that teachers could utilize voluntarily on their own time. These supports aim to deepen teacher learning by offering additional resources and materials teachers can use to address their professional needs or interests. Teachers were introduced to the online supports during the learning institutes that were offered at each school. EL Education hosts interactive webinars on a multitude of curricular, instructional practice, or general education topic areas, where teachers could engage with newly presented information and have opportunities to ask questions and engage in a dialogue with presenters and other webinar participants. Examples of the topics that webinars focused on included "Management in the Active Classroom: Concrete Practices that Maximize Student Learning," "Models, Critique and Descriptive Feedback," and "Word Work: Helping Students Succeed through Effective Vocabulary Instruction." EL Education offered up to three webinars and treatment schools were allowed to choose which webinars topics they wanted.

TPP online supports also included Professional Learning Packs (PLPs), which are web-based libraries of PD materials dedicated to a particular theme. These PLPs are open educational resources located on EL Education's website that teachers could access and use as needed to support their professional development. They contain selected texts and videos that TPP teachers could use to refine their practices while also applying what they were learning from the institutes, webinars, and TPP curriculum. Examples of PLPs included "Standards and Shifts: Understanding the Common Core," "Teaching Reading," and "Student-Engaged Assessment." Other online supports included resources available on EL Education's website for teachers to use as part of their independent study, including models of student work that were aligned to CCSS as well as links to books, videos, toolkits, and guides that covered exemplary pedagogical practices as well as curriculum and classroom management strategies.

2. Second year of TPP implementation for selected schools

The structure of TPP in the second year of implementation that took place among selected schools remained largely the same as what was implemented in the first year, in that participating

schools continued to use the CCSS-aligned ELA curriculum, and engage in the three components of PD support (institutes, coaching, and online supports). However, there were several differences in the magnitude or content of the institutes and coaching that were provided in the second year relative to those offered in the first year.

- **Institutes:** Whereas a common set of five institutes were provided to all participating schools in the first year of TPP, the second year of TPP included only one institute. The topic focus and audience of the institute was determined collaboratively between the TPP coach and the staff at each school. These institutes were connected to the schools' goals and helped address their particular needs. For example, the Using Module Assessment Data institute provided in the second year of implementation of TPP helped teachers use the curriculum's assessments to gather information on students' progress in general and against the ELA CCSS at each grade level.

- **Coaching:** There were relatively fewer coaching days in the second year of implementation of TPP, with coaches being on-site for 16 days and offering 2 days of video-based coaching (as opposed to the 21 days of on-site and 3 days of video-based coaching in the first year). The amount of time dedicated to strategic planning at the schools remained unchanged in the second year (1.5 days). The coaching provided in the second year had an emphasis on developing teacher and school leaders and building capacity at the school. The aim of this second-year focus was to deepen the progress of the first year so that the curriculum and instructional shifts that teachers and schools learned during their participation in TPP could continue after the formal intervention period ended. With coaching in the first year focused largely on fostering teachers' mastery of the TPP curriculum and facilitating a teacher-centered approach toward enhancing teachers' practices in the classroom, coaching in the second year shifted more toward a student-centered approach. For example, coaches worked with teachers to use student assessment data from the curriculum modules to make instructional decisions based on their students' needs.

B. Evaluating TPP implementation fidelity

The implementation evaluation of the TPP program is guided by the overarching research question: "Is there evidence of intervention fidelity?" To help address this research question, we assess the extent to which TPP's key components were implemented with fidelity through our assessment of data derived from a variety of sources. In this section, we describe these data sources and how they were used to examine TPP's implementation fidelity. We then describe the implementation fidelity of the TPP curriculum, how we developed measures for determining the implementation fidelity of the PD components, and the implementation findings for those components.

1. Data sources

The implementation evaluation included data that came from multiple sources: TPP coach surveys and teacher TPP participation forms, teacher surveys, PD exit surveys, and classroom

observations. The data sources and specific sample information are discussed in more detail below.

Coach surveys and teacher TPP participation forms. There were a total of 15 coaches who supported the 35 treatment schools that received one year of TPP, along with 4 coaches who worked with the 10 treatment schools that received two years of TPP. Coaches were asked to complete a separate survey for each school they worked with, and there was a 100 percent survey response rate for the one-year TPP schools and a 90 percent response rate for the two-year TPP schools. Open-ended questions were used to determine if there were any school-level instances of the TPP curriculum not being implemented and to identify patterns related to the implementation of the PD supports.

Coaches completed the teacher TPP participation form for each ELA teacher within treatment schools to determine the extent of their participation in the PD supports, particularly the institutes and coaching components. The participation form was considered our primary source of data for these PD supports because we had information on all ELA teachers in the sample—a form was completed for all 162 one-year TPP ELA teachers in the 35 treatment schools and 34 two-year TPP teachers in the 10 treatment schools. This includes the 48 one-year novice ELA teachers in 29 treatment schools and the 12 two-year novice ELA teachers in 7 treatment schools (see discussion in Chapter II.A).

Teacher surveys. The teacher survey was used to assess implementation of the online supports PD component because this was a TPP component that teachers utilized voluntarily and on their own time and TPP coaches may not be fully knowledgeable about teachers' access to these supports. Teachers were asked whether they accessed online supports, such as webinars, PLPs, online facilitated conversations, online PD materials, and the online community of practice.

There were a total of 135 one-year TPP ELA teachers who completed the professional development section on at least one survey across the fall or spring (83 percent response rate). This includes 46 one-year novice ELA teachers (a 96 percent response rate).[12] At least one teacher, including novices, from each of the schools completed this section of the survey. There were a total of 29 two-year TPP ELA teachers who completed the professional development section on at least one survey across the fall or spring (a 85 percent response rate), which includes 34 two-year novice ELA teachers (a 100 percent response rate).[13] At least one teacher from 9 out of 10 two-year TPP schools completed this section of the survey; at least one novice teacher completed at least one survey at all the 7 two-year TPP schools that had novice teachers. The five two-year teachers (none were novices) who did not provide any survey data on the online support PD component were from 5 schools.

PD exit surveys. The PD exit survey, which was only administered in the one-year TPP schools, was used to provide contextual information on participants' experiences with the TPP institutes.

[12] In one-year TPP schools, there was a 69 percent response rate in the fall and 78 percent response rate in the spring.

[13] In two-year TPP schools, there was a 79 percent response rate in the fall and 77 percent response rate in the spring.

Respondents reported on how useful they thought the content of the institute was and how they intended to use the information in their classrooms or schools. Exit surveys were provided to participants at each of the five institutes that were offered to the one-year TPP schools. A total of 410 PD exit surveys were collected across all study schools; 126 of those responses were from novice ELA teachers. However, an important limitation to these data is that the surveys were completed anonymously and an individual could complete the survey up to five times (one for each institute), and any school staff that attended the institute could complete a survey (including non-ELA staff and school administrators). As a result, response rates cannot be determined for this data source.

Classroom observations. Trained observers visited all study schools, treatment and control, to assess the CCSS-aligned instructional practices teachers used in the classrooms. When visiting treatment schools, observers recorded which TPP curriculum module was being implemented in the observed classroom, and this information was used to complement coaches' reports on the extent to which the TPP curriculum was implemented in study schools. In one-year TPP schools, we observed 62 percent of all ELA teachers, which includes 96 percent of the novice ELA teachers. In two-year TPP schools, we observed 91 percent of all ELA teachers, which includes 100 percent of the novice ELA teachers

2. TPP ELA curriculum implementation fidelity findings

According to coaches' reports and the classroom observations, the TPP CCSS-aligned ELA curriculum was implemented in all treatment study schools that participated in the first and second year of TPP. This aligned to the agreement set with TPP schools that they implement TPP's CCSS-aligned ELA curriculum in ELA classrooms for a minimum of 45 minutes each day.

In their surveys, coaches discussed both the successes and challenges schools and teachers experienced with implementing the curriculum in the classrooms. Coaches reported how teachers strove to adhere to the lesson plans in the curriculum modules, such as introducing the learning targets and using the workbooks, texts, and assessments included in the modules. Yet, there were also implementation issues that arose that produced challenges for delivering the curriculum as intended. For example, as teachers were working to familiarize themselves with the new curriculum and materials, they experienced difficulties with pacing and delivering a given lesson plan during the class period. In these instances, teachers would prioritize different parts of the lesson and at times would not be able to deliver all aspects of a particular curriculum module.

3. Developing implementation fidelity measures for the PD components

We developed measures to assess fidelity of implementation at study schools through a two-step process, based on whether teachers engaged in each of TPP's three PD components. Below we describe the implementation fidelity measures for each of the PD components. Specifically, we describe how we created

Key finding: The TPP ELA curriculum was implemented in all treatment schools.

teacher-level participation measures on each component and how we used these to develop school-level fidelity measures to gauge whether each component was implemented with fidelity.

First, we created the following teacher-level participation measures for each PD component.

Institute attendance. The institutes are an essential component for teachers' implementation of the TPP curriculum in that they aim to provide teachers with a strong understanding of what CCSS shifts are and how the curriculum aligns to those shifts. It also offers teachers opportunities to hone their instructional practices to carry out the curriculum in their classrooms as it was designed. For teachers' first year of TPP, five institutes were offered to schools. We classified teachers' participation in these institutes as low (did not attend any institutes), medium (attended at least one institute), and high (attended at least two institutes). Because only one institute was provided at each school in the second year of TPP, the teacher-level participation measure was adjusted to low (not attending any institute days) and high (attending one institute day).

Coaching participation. TPP coaching offered TPP teachers ongoing professional supports throughout the school year, focusing on providing teachers with regular feedback on their classroom instructional practices and supporting their implementation of the CCSS-aligned curriculum. At the teacher level, we classified teachers' participation as low (met with their coach less than 5 times during the school year), medium (met with their coach 5 to 9 times during the school year), or high (met with their coach at least 10 times during the school year). These teacher-level participation classifications were applied to both years of TPP implementation.

Online supports utilization. The online supports offered to TPP teachers consisted of a variety of materials and informational resources teachers could consult in their preparation or during their ongoing implementation of the TPP curriculum. At the teacher level, we classified teachers' participation as low (did not access any of the online resources), medium (accessed at least one of the online resources), or high (accessed two or more of the online resources). These teacher-level participation scores were applied to both years of TPP implementation.

Second, using these three sets of teacher-level participation measures, we then created school-level fidelity measures based on the percentages of teachers within the school that had high teacher-level participation. For the institutes and coaching components, we classified schools as low (no ELA teachers have high participation scores), medium (less than 50 percent of ELA teachers have high participation scores), or high (50 percent or more of ELA teachers have high participation scores). For the online supports indicator, we classified schools as low (none of the ELA teachers had medium or high teacher-level participation score), medium (less than 50 percent of ELA teachers had medium or high teacher-level participation scores), or high (50 percent or more of ELA teachers had medium or high teacher-level participation scores).

We created separate school-level fidelity measures when only including novice teachers in the sample of teachers at a given school and when including all ELA teachers (novices and experienced) in the sample of teachers at a given school. This allowed us to determine if school-

level implementation occurred among novices (referred to as school-level novice teacher fidelity) and schoolwide (referred to as school-level all teacher fidelity). At the study level, we then classified whether each of these TPP PD components was implemented with fidelity if at least 50 percent of schools had high fidelity scores, assessing TPP's implementation among novices and all ELA teachers separately, and in the first year and second year separately.

4. TPP professional development implementation fidelity findings

Using the teacher participation and school-level fidelity measures described above, we assessed the extent to which each of the TPP PD components was highly implemented at the study level. We complement these analyses with data from TPP coach surveys and PD exit surveys to provide more descriptive information on how TPP participants perceived these PD supports and coaches' experiences with implementing TPP in study schools.

Key finding: School-level implementation fidelity of the PD components was high in the first and second years of TPP among the novice ELA teachers at these schools, except for novice ELA teachers accessing online supports during the second year.

Attending institutes. We assessed the school-level implementation fidelity of attending institutes among the 29 TPP schools that had at least one novice teacher in their first year of the study, based on coaches' reports using the teacher TPP participation form. We found that 76 percent of schools demonstrated high school-level novice teacher fidelity (Table IV.1). Within each of those high-fidelity schools, at least half of the novice teachers participated in two or more institutes. The results were relatively similar when looking at implementation fidelity among all teachers (novices and experienced) at all 35 schools in their first year of TPP. More than two-thirds of all TPP schools in their first year had high school-level fidelity with regards to all ELA teachers participating in two or more institutes. Overall, the institute's component of TPP's PD was implemented with fidelity in the first year of TPP, among both novice and all teachers in the study schools in that more than half of the schools had high implementation. However, there was some school-level variation, as less than one-quarter of study schools had low implementation.

Institutes continued to be implemented with fidelity among the study schools receiving a second year of TPP among both novice and all ELA teachers in those schools. When looking at implementation fidelity among the TPP schools that had at least one novice ELA teacher in our eligible sample (n = 7), 57 percent had high school-level novice teacher fidelity. The implementation fidelity among all ELA teachers at TPP schools in their second year of TPP followed a similar pattern. However, unlike implementation in the first year of TPP, approximately 40 percent of schools had low implementation in the second year of TPP.

According to the teacher exit surveys, novice teachers felt that these institutes provided valuable learning experiences. Nearly all novice teacher responses (94 percent) found the institutes to be useful or very useful, with the remaining 6 percent reporting that the institutes were slightly useful. Among all institute attendees (including all teachers and school leaders), 89 percent reported that the institutes were useful or very useful, 10 percent found them to be slightly useful, and 1 percent stated that they were not useful. This pattern, among novice teachers and all

attendees, was similar across nearly all schools and the different institutes that were provided. For example, 85 percent of all attendees found that the Using Data to Sharpen Implementation institute was useful. They explained that it helped them use data more accurately and look for trends in the data to help their students. Other attendees explained that this institute provided them with useful information on how to implement rubrics in their classrooms and how to review rubrics with their students.

Table IV.1. The percentage of TPP schools with low, medium, or high implementation fidelity of institutes for novice and all teacher samples, by year of TPP implementation

TPP component	Percentage of TPP schools with low implementation	Percentage of TPP schools with medium implementation	Percentage of TPP schools with high implementation	Was TPP component implemented with fidelity?
Attending TPP institutes				
Year 1				
Novices (*n* = 29)	24.1	0.0	75.9	Yes
All teachers (*n* = 35)	17.1	14.4	68.6	Yes
Year 2				
Novices (*n* = 7)	42.9	0.0	57.1	Yes
All teachers (*n* = 10)	40.0	0.0	60.0	Yes

Note: A TPP component is classified as being implemented with fidelity if at least 50 percent of schools have a "high" implementation score.

Coaches also reported that attending institutes facilitated the implementation of other TPP components in study schools. They stated that the institutes helped teachers better understand and implement the TPP curriculum, especially when the institutes were held before teachers began implementing certain curriculum modules. While there were sometimes difficulties with scheduling institutes around teachers' availability and other school calendar events in some schools, institutes were more readily implemented with teachers when school leadership would set aside teachers' time for the institutes or when they demonstrated their support for the institutes by encouraging teacher attendance at them.

Working with a TPP coach. When looking at the implementation fidelity among novice ELA teachers based on coaches' reports using the teacher TPP participation form, 86 percent of the schools in their first year of TPP had high school-level fidelity with regards to working with a TPP coach (see Table IV.2). In other words, 86 percent of these TPP schools had at least half of their novice ELA teachers participate in coaching at least 10 times during the school year. High school-level all teacher fidelity was also exhibited in 57 percent of all TPP schools, as over half of all ELA teachers participated in coaching at least 10 times during the school year.

There was also high implementation fidelity with TPP coaching among schools in their second year of TPP, for both novice ELA teachers and all ELA teachers. All second-year schools that had at least one novice ELA teacher were classified as having high school-level novice teacher fidelity, with at least half of novice ELA teachers at these schools meeting with a TPP coach ten

times or more during the year. Using a sample of all ELA teachers, 60 percent of schools in their second year of TPP had high school-level all teacher fidelity on the coaching component of TPP. Overall, the coaching PD component of TPP was implemented with fidelity in schools among both novices and all teachers and across schools' participation in years one and two of TPP, as more than 50 percent of the schools had high implementation. However, there was some variation, especially with the one-year and two-year samples of schools among all ELA teachers, as approximately 40 percent of the schools did not have high implementation.

Table IV.2. The percentage of TPP schools with low, medium, or high implementation fidelity of coaching for novice and all teacher samples by year of TPP implementation

TPP component	Percentage of TPP schools with low implementation	Percentage of TPP schools with medium implementation	Percentage of TPP schools with high implementation	Was TPP component implemented with fidelity?
Working with a TPP coach				
Year 1				
Novices (*n* = 29)	13.8	0.0	86.2	Yes
All teachers (*n* = 35)	11.4	31.4	57.1	Yes
Year 2				
Novices (*n* = 7)	0.0	0.0	100.0	Yes
All teachers (*n* = 10)	20.0	20.0	60.0	Yes

Note: A TPP component is classified as being implemented with fidelity if at least 50 percent of schools have a "high" implementation score.

Coaches reported on a variety of successes and challenges with regards to implementing the coaching component of TPP in study schools. Coaches felt that teachers were responsive to the coaching they received, such as allowing coaches to co-teach part of a lesson and being open to feedback from coaches; this allowed teachers to enhance the strategies they were seeking to implement in their classrooms. The planning sessions with teachers and their grade-level teams, along with working with teachers' school instructional coaches, fostered cohesion among the teaching teams and provided opportunities to address any reservations they may have had. However, coaches encountered difficulties with providing coaching, particularly when on-site, because of the structural barriers or competing priorities that school staff faced. For example, teachers were not always able to access coaches while they were on-site during the school day because coverage could not be found for their classrooms; as a result, they need to reach out to coaches outside of regular school hours. Moreover, while school leaders generally supported TPP and the supports provided to their teachers, the school leaders themselves were not always available or able to join strategic planning sessions with coaches because of time constraints and prioritizing their management responsibilities.

Accessing online supports. Based on those who responded to the teacher survey at least one time (fall or spring or both) during their school year, we assessed the extent to which ELA teachers accessed TPP online supports and whether this PD component of TPP was implemented with fidelity (see Table IV.3). In the first year of TPP, 62 percent of schools had at least half of

their novice ELA teachers access online supports at least one time during the school year. When looking at the implementation fidelity among all teachers (novices and experienced) at TPP schools in their first year, 43 percent had at least half of all their ELA teachers access the online supports at least one time during the school year. So, while there was implementation fidelity of the online supports in schools in their first year among novice ELA teachers, this was not the case when examining all ELA teachers' accessing online supports.

In schools' second year of TPP, teachers' survey responses indicate that the online supports component of TPP was not implemented with fidelity. When looking at the implementation fidelity among novice teachers, 43 percent of these TPP schools had at least half of its novice ELA teachers access online supports one time or more during the school year, but more than half of the schools had novice teachers that did not access the supports at all during the second year of TPP. A similar pattern emerged when assessing the school-level implementation accessing online supports when considering the behavior of all ELA teachers at these schools.

Table IV.3. The percentage of TPP schools with low, medium, or high implementation fidelity of online supports for novice and all teacher samples by year of TPP implementation

TPP component	Percentage of TPP schools with low implementation	Percentage of TPP schools with medium implementation	Percentage of TPP schools with high implementation	Was TPP component implemented with fidelity?
Accessing online supports				
Year 1				
Novices (*n* = 29)	34.5	3.4	62.1	Yes
All teachers (*n* = 35)	20.0	37.1	42.9	No
Year 2				
Novices (*n* = 7)	57.1	0.0	42.9	No
All teachers (*n* = 9)	55.6	11.1	33.3	No

Note: A TPP component is classified as being implemented with fidelity if at least 50 percent of schools have a "high" implementation score. There was one TPP school in year 2 where none of the experienced teachers at the school had survey information on their use of online supports (there were no eligible novice teachers at this school).

Unlike the data source for the institutes and coaching PD components of TPP, which was the teacher TPP participation form that was completed by coaches for all teachers in the study, the teacher survey used to assess the online supports component had missing data from some teachers. Because some teachers did not complete the teacher survey in both the fall and spring in a school year, or did not complete a survey at all, there is potential for this missing data to bias the results of whether the online supports component was implemented with high fidelity in the study schools.[14] As a result, we created lower bound and upper bound assumptions about how

[14] There were 27 teachers from 13 schools (including 2 novices from 2 schools) who did not provide any survey data on the online supports PD component. The 5 two-year teachers (none were novices) who did not provide any survey data on the online supports PD component were from 5 schools.

this missing data might affect the results. More specifically, if teachers were missing data from the fall and/or spring time point, the lower bound assumption was that the teachers did not access online supports at all, whereas the upper bound assumption assumes that the teachers did. We then recalculated teacher-level and school-level fidelity on the implementation of the online supports, with the results presented in Table IV.4, to see how this missing data could impact our characterization of implementation fidelity.

Making the lower bound assumption (that teachers with missing data did not access the online supports) did not change the school-level implementation fidelity pattern among schools in their first and second year of TPP and with considering novice and all ELA teacher samples. For example, access to online supports had implementation fidelity in schools in their first year of TPP among novice ELA teachers using this lower bound assumption, which provides greater confidence in the results presented in Table IV.3 for this sample of schools and novice teachers.

Making the higher bound assumption (that teachers with missing survey data did access online supports) did change the school-level implementation fidelity pattern. Using this assumption, online supports were implemented with fidelity in schools during their first and second year of TPP, among both the novice and all ELA teacher samples. While it is possible that the online supports were implemented with fidelity among study schools, and the level of missing survey data inhibited this pattern from emerging, the findings discussed in Table IV.4 more closely parallel the lower bound assumption for assessing the implementation of this PD component. Moreover, several coaches in their surveys cited online supports, such as webinars, as one TPP component that was not being implemented well in their schools. For example, they reported that online webinars were not heavily attended by teachers or school leaders at their schools because they were held after school and their attendance was not mandatory.

Overall assessment of intervention fidelity of the study. All treatment schools in the study implemented the TPP CCSS-aligned ELA curriculum. School-level implementation fidelity of the PD components was high in the first and second years of TPP among the novice ELA teachers at these schools, with one exception: there was low school-level fidelity of novice ELA teachers accessing online supports during the second year. School-level implementation fidelity among all ELA teachers followed a similar pattern. School-level implementation fidelity was high among all ELA teachers in the study schools in their first and second year of TPP in several dimensions: schools implementing the ELA curriculum, teachers attending institutes, and teachers meeting with coaches. However, there was low implementation fidelity of accessing online supports in either year of the study for schools when looking at all ELA teachers in those schools.

Table IV.4. The percentage of TPP schools with low, medium, or high implementation fidelity of online supports for novice and all teacher samples by year of TPP implementation using lower and upper bound assumptions

TPP component	Percentage of TPP schools with low implementation	Percentage of TPP schools with medium implementation	Percentage of TPP schools with high implementation	Was TPP component implemented with fidelity?
Accessing online supports (lower bound assumptions)				
Year 1				
Novices (*n* = 29)	34.5	3.4	62.1	Yes
All teachers (*n* = 35)	20.0	40.0	40.0	No
Year 2				
Novices (*n* = 7)	57.1	0.0	42.9	No
All teachers (*n* = 10)	60.0	20.0	20.0	No
Accessing online supports (upper bound assumptions)				
Year 1				
Novices (*n* = 29)	20.7	0.0	79.3	Yes
All teachers (*n* = 35)	5.7	17.1	77.1	Yes
Year 2				
Novices (*n* = 7)	28.6	0.0	71.4	Yes
All teachers (*n* = 10)	10.0	20.0	70.0	Yes

Note: A TPP component is classified as being implemented with fidelity if at least 50 percent of schools have "high" implementation scores.

This page has been left blank for double-sided copying.

V. IMPACTS OF TPP ON TEACHERS' INSTRUCTIONAL PRACTICES

TPP aims to prepare teachers to engage students in high-quality instruction aligned with CCSS. To assess whether TPP achieved this goal, we examine the impact of TPP on teachers' ELA instructional practices, including CCSS-aligned instructional practices of novice and all ELA teachers. The study team assessed one-year and two-year impacts of TPP on teacher ELA instructional practice outcomes based on classroom observations and teachers' self-reports from surveys (research questions T1 through T4 shown in Table I.1). For this report, the teacher ELA instructional practice outcomes are defined as broader outcomes that combine several instructional practices.

To present a more comprehensive understanding of the impacts of TPP on teachers' ELA instructional practices, we discuss these findings alongside earlier study findings (Choi et al. 2017, Choi et al. 2018). The aim of the earlier research was to provide early insights into the effects of TPP to highlight areas to inform implementation successes and areas for improvement. For this reason, the earlier research focused on TPP's effects on individual ELA instructional practices to capture nuanced and specific practices teachers utilized; this report, in contrast, focuses on the combination of several instructional practices. Specific instructional practices were statistically grouped into constructs if they were similar; for example, teachers encouraging students to cite evidence from the text in discussions, supporting students to find evidence from a text to support their responses, and asking students to use text evidence to write an argumentative essay are related to each other and grouping them together provides a broader understanding of TPP's impacts on teachers' instruction related to students' use of evidence from texts.

This chapter begins with a description of how teacher instructional practice constructs were developed for this report. We next provide information on the sample of teachers and the schools they were in for the one- and two-year analyses, as well as our analytic approach to estimating impacts and the ways the current approach differed from that of the earlier analyses of teachers' instructional practices. We then present the findings from the current approach alongside findings from earlier analyses. We conclude this chapter with key considerations for interpretation.

A. Developing teacher ELA instructional practice constructs

Teacher instructional practices are complex and involve a number of behaviors, strategies, and tasks. For example, the practice of supporting students' higher-order thinking could involve specific activities of teachers asking students probing questions to encourage students to think at a deeper level, teachers encouraging students to critique a topic they are discussing, and teachers supporting students to make predictions about what will happen next in a text they are using. Because of the complex nature of instructional practices, a robust way to capture a practice is by creating instructional practice constructs using statistical methods—exploratory factor analysis and Bayesian confirmatory factor analysis—to combine multiple observation or survey items that focus on fine-grained and specific teacher activities.

We developed constructs by first creating a conceptual framework that relates the survey and observation items to teacher instructional practices. This conceptual framework focused on topics within three main areas of instructional practices: (1) general instruction, (2) reading and writing instruction, and (3) classroom management and environment. These three main areas of instructional practices were identified by our expert review panel, which included experienced ELA teachers, ELA curriculum and professional development developers, and professors and researchers of ELA education who have worked with teacher candidates and on ELA education studies. Within these main areas, the study team conceptually identified topics, such as higher-order thinking, and conducted exploratory factor analysis to determine the optimal number of constructs within each instructional practice topic. Finally, the study team used Bayesian confirmatory factor analysis to finalize our constructs (see Table V.1), which were each based on a number of classroom observation or survey items. We used 105 items to develop 31 constructs across 16 topics in the three main areas using data from year 1. The same constructs were used for both years of the study. Model fit statistics and internal-consistency reliability suggest the constructs were a good fit to the data for year 1 and an adequate fit for year 2. Below is a description of the topic areas and constructs organized by the three main areas (see Appendix Table E.4 for additional details, including the items included in each construct).

General instruction. The general instruction area captures instructional practices that support student learning. The topics in this area focused on whether teachers had well-structured lessons; connected students' learning to their prior knowledge, personal lives, or the real world; focused on instruction supporting students' higher-order thinking and content knowledge development; encouraged students' participation in class and discussions; supported students' responsibility for their own learning. A description of the 16 constructs within the 9 topics in this area is presented in Table V.1.

Table V.1. General instruction practice area: topics, constructs, and construct descriptions

Topic	Construct (Number of items)	Construct description
Connected lessons	Connected lessons (2)	Activities within a lesson were connected to one another and the lesson included a summary at the end to synthesize learning.
	Stated learning purpose (2)	The teacher or students stated the purpose of instruction out loud and the teacher referred back to the purpose of instruction during the lesson.
Connections to world	Learning connected to personal experiences or real world (3)	The students and teacher connected students' learning—including through reading activities and with academic vocabulary instruction—to the real world or a personal example.
Content knowledge	Content knowledge development (3)	The extent to which and the amount of class time the teacher focused on building students' content knowledge and the degree to which content knowledge was connected to the main point of the lesson.
	Demonstrate content knowledge in writing (2)	When grading, evaluating, or providing feedback to students' writing, the degree to which teachers felt it important that students demonstrate their content knowledge.
Higher-order thinking	Assessment of higher-order thinking (4)	The extent to which teachers felt the development of ideas and use of evidence to support ideas, and quality of writing were important when grading or evaluating students' writing, and when providing students with feedback on their writing.
	Higher-order thinking and reading (7)	The frequency with which teachers focused on the following when students read texts in class: text analysis, analyzing differences and similarities in perspectives between multiple texts on the same topic, critiquing or evaluating arguments and specific claims in context, determining central idea(s) or theme(s) of the text, determining the author's point of view or purpose, integrating or comparing and contrasting information in different media or formats, and integrating or interpreting information across sections of the text.
	Observed higher order thinking (7)	The extent to which the teacher asked higher-order thinking questions either verbally or in writing, reinforced higher-order thinking in class, supported students' analysis of text, and asked questions to encourage students' to think at a higher level, including probing questions or follow-up questions.
Multimedia use	Online writing (2)	The frequency with which teachers asked students to create videos or webcasts and the frequency with which teachers asked students to write for social networking, blogs, or wikis.
	Use of multimedia in instruction (3)	Whether the teacher used the following in the class period to provide instruction: (1) chalk board, smart board, projector, or chart paper; (2) games, role play, drama; and (3) films, videos, music, and art.
Students' prior knowledge	Prior knowledge (2)	The frequency with which teachers asked students to relate what they read to their own experience or to something they have learned before in the past two weeks, and how often teachers focus on having students relate the story or literary work, its characters, and/or its themes to something they have read before when students read texts in class.
Student independence	Student independence (3)	The frequency with which teachers encouraged students to engage in the following behaviors in the past two weeks: to ask for more challenging work, to check their own progress against learning targets, to give input in setting the learning targets, and to help develop rubrics or evaluation criteria against learning targets.
Student participation	Collaborative discussion practices (4)	The frequency with which teachers' supported students' discussion practices to respond to peer comments and/or build on each other's thinking, whether the teacher used student pair or group work in the class period, whether the teacher asked other students to help answer the question in discussions, and the degree of thoughtfulness of students' responses during discussion.

Table V.1 (*continued*)

Topic	Construct (Number of items)	Construct description
	Student engagement in class (2)	The approximate proportion of students who struggled to stay on task during the class period and whether the majority of students were on task for almost all the class period.
	Student participation in discussion (2)	The extent to which students participated in class and the extent to which students spoke most in class.
Student responsibility for learning	Student responsibility for their learning (3)	The frequency with which students were supported to provide feedback on each other's work, share their work with their peers, and work with other students in the past two weeks.

Reading and writing instruction. The reading and writing instruction area was defined to capture instructional practices that support students' engagement with texts and reading and writing practices that are the foci of CCSS. There are five topics within the reading and writing instruction area that focus on whether teachers engaged in certain reading and writing instructional practices: integrating academic vocabulary; having students engage in multiple types of writing and write for multiple purposes; having students engage with texts by reading, writing, and/or speaking about texts; using evidence from text to support their ideas; and focusing on developing students' writing conventions. A description of the 11 constructs within the five topics in this area is presented in Table V.2.

Table V.2. Reading and writing practice area: instruction topics, constructs, and construct descriptions

Topic	Construct (Number of items)	Construct description
Academic vocabulary	Academic vocabulary (2)	Degree to which the teacher focused on academic vocabulary instruction and whether students engaged in reading activities focused on determining word meanings from texts.
Multiple types and purposes of writing	Engagement in multiple types of writing (2)	Number of writing purposes, genres, and modes of writing teachers asked students to engage in.
	Writing for multiple purposes and audiences (2)	Number of audiences and types of writing (for example, writing to inform or writing to argue).
Reading, writing, and/or speaking about texts	Close reading and writing that demonstrates understanding of text (5)	Whether teachers asked students to engage in close reading or to review the text to find evidence to support a viewpoint. Whether teachers asked students to write to demonstrate their understanding of the text and to cite sources or evidence from texts.
	Reading, writing, and/or speaking about texts in past two weeks (4)	Frequency with which teachers supported students as they discussed texts they read with partners or a small group of students, produced extended writing by incorporating key details from texts they read, shared their ideas about and/or understanding the texts they read with the whole class, and wrote about texts they read in the past two weeks.
	Reading, writing, and/or speaking about texts in typical week (5)	Frequency with which teachers asked students to do a group activity or project about what they have read, to explain or support their understanding of what they have read, to write about something they have read, to read books they have chosen themselves, and to describe the style or structure of the text they have read in a given week.
Use of text evidence	Assessment of text evidence use (2)	The importance of the accuracy or appropriateness of students' references to texts when teachers graded or evaluated students' writing and provided students with feedback on their writing.
	Self-reported text evidence use (2)	In the past two weeks, the frequency with which teachers asked students to write citing evidence or information from text(s) and to write thinking tasks, such as graphic organizers, that capture students' thoughts on the text and relevant evidence.
	Text evidence use and argument writing (4)	Whether the teacher asked students to cite evidence from the text to support their responses either verbally or in writing. Whether teachers supported students' argument in writing and writing activities that involved citing sources or evidence with or without the teacher's help.
Writing conventions	Feedback on writing conventions (6)	The importance of writing conventions and effective use of language when teachers graded or evaluated students' writing or provided feedback on students' writing. The importance of the organization of ideas when teachers graded or evaluated students' writing.
	Use of writing conventions (2)	Teachers asked students to write to practice writing conventions and students engaged in writing activities that involved writing conventions.

Classroom management and environment. The classroom management and environment area was defined to capture the practices teachers engaged in to support classroom management and create a positive classroom environment. The area had two topics: classroom climate and classroom management. A description of the four constructs within the two topics in this area is presented in Table V.3.

Table V.3. Classroom management and environment practice area: topics, constructs, and construct descriptions

Topic	Construct (Number	Construct description
Classroom climate	Classroom climate (6)	Degree to which the teacher and students demonstrated respect for one another. Degree to which the teacher was warm, supportive, and passionate about what they were teaching. Proportion of students who were receptive to the teacher's instructions and feedback and were motivated. Degree to which teacher successfully encouraged students to interact with one another.
Classroom management	Classroom disruption (reverse) (3)	The extent to which students disrupted class for more than 5 minutes and more than 10 minutes. The extent to which transitions between activities took more than 5 minutes. The items in this construct were reverse coded to have a similar scale as other constructs (that is, higher scores are more positive instructional practices than lower scores).
	Observed classroom management (3)	Degree to which teachers' attempts to redirect student disruptions were effective and teachers enforced behavioral rules in the classroom. Whether teachers' redirection of student disruption took away from learning time.
	Self-reported classroom management (6)	The extent to which teachers emphasized positive character traits and values in the classroom, managed their class well, acknowledged positive student behavior, redirected students back to the topic when they get off-task, required that students in the classroom follow the rules at all times, and set clear expectations for student behavior.

Overall ELA instructional practices. The overall ELA instructional practices area captures all the ELA instructional practices defined as important to supporting students' ELA learning. The overall ELA instructional practices include the 16 topics in all three of the aforementioned areas: (1) general instructional practices, (2) reading and writing instruction, and (3) classroom management and environment.

After instructional practice constructs were developed, we calculated teacher scores for each construct using complete case analyses, which means that only the teachers who had responses for each item in a given construct were included in the analyses. The overall ELA instructional practice scores were also calculated using complete case analyses, which means that only the teachers who had a score for each construct had overall scores calculated. Appendix E provides more information on the specific processes we conducted to develop these constructs, information on which questions were included for each construct, model fit information, and more information on how we calculated the composite scores.

B. Description of teacher samples

To provide insight into the characteristics of the schools where teachers taught and the students they served, as well as to test whether there were any differences between treatment and control teachers' schools and students, in this section we present teachers' school and student characteristics for teachers who contributed observation or survey data. Additionally, to provide information on the extent to which teachers included in the analyses are representative of the sample of teachers, we present response rate and attrition information. This section first discusses the sample of teachers in the one-year study and then the sample of teachers in the two-year study.

1. Sample of teachers in the one-year study and description of school and students

Teachers who provided observation and survey data in the one-year study were from 69 schools (34 control and 35 treatment) across Cohorts 2 and 3.[15] These teachers were part of the schools that were randomly assigned to use TPP (treatment) or to continue business as usual (control) for the 2015–2016 (Cohort 2) or 2016–2017 (Cohort 3) school year.

The majority of the schools were noncharter schools, had between 36 and 37 full-time teachers on average, and served approximately 560 students, more than 70 percent of whom qualified for free or reduced-price lunch (see Table V.4).

Table V.4. Comparison of school characteristics at baseline for the one-year impact analysis

	Treatment		Control		t-statistic	Effect size
	Mean	SD	Mean	SD		
Whether charter school (percent)	2.90	0.17	0.00	0.00	0.97	0.24
Number of full-time teachers	37	15	36	11	0.48	0.12
Percent students eligible for free or reduced-price lunch	73.40	0.24	70.90	0.22	0.45	0.11
Number of students enrolled	556	196	568	233	-0.21	0.05

Note: These figures are based on the 66 (32 treatment and 34 control) schools for which the district provided school-level demographic data. There were no statistically significant differences between control and treatment schools on these characteristics.

In these schools, there were 361 ELA teachers (199 control and 162 treatment) who were eligible to be part of the study. Of these 361 teachers, 113 were novice teachers (65 control and 48 treatment). Among teachers who contributed observation and/or survey data, teachers taught students who were predominantly white non-Hispanic or black non-Hispanic and half of whom were male (see Table V.5). A small percentage—less than 15 percent—were classified as English-language learners or as needing special education services. We assessed baseline equivalence for novice and all teachers on their students' standardized reading pre-test scores,

[15]Numbers of schools for teachers who were included in the one-year analyses are included in Appendix V.

standardized math pre-test scores, proportion male, race (white non-Hispanic, black non-Hispanic, and other), Hispanic ethnicity, free and reduced-price lunch eligibility, English-language learner status, and special education status. Statistically significant differences and effect sizes greater than 0.05 standard deviations indicated the presence of differences between the treatment and control groups. Among the novice teachers, there were differences in the proportion of special education students in the treatment and control classrooms. For all teachers, there were differences in the proportions of white non-Hispanic and black non-Hispanic students in the treatment and control classrooms. Descriptions of the characteristics of students of teachers who were included in our analyses are presented in Appendix Tables E.10 and E.11.

Table V.5. Comparison of characteristics of teachers' students at baseline for the one-year impact analysis

	Treatment			Control			t-statistic	Effect size
	Mean	SD	N	Mean	SD	N		
Novice teachers								
Average student standardized ELA pre-test score	-0.60	0.48	43	-0.52	0.44	49	0.89	0.19
Average student standardized math pre-test score	-0.57	0.48	43	-0.53	0.47	49	0.40	0.08
Percent students male	53.1	0.08	43	53.0	0.07	49	-0.07	0.02
Student race and ethnicity								
Percent white non-Hispanic	34.7	0.35	43	39.8	0.32	49	0.73	0.15
Percent black non-Hispanic	41.8	0.38	43	35.8	0.30	49	-0.85	0.18
Percent Hispanic	17.1	0.23	43	17.0	0.17	49	-0.04	0.01
Percent other race	6.3	0.06	43	7.5	0.08	49	0.84	0.18
Percent students eligible for free or reduced-price lunch	76.7	0.24	40	73.6	0.23	48	-0.61	0.13
Percent students English-language learners	6.7	0.10	43	6.3	0.08	49	-0.19	0.04
Percent students special education	14.2	0.09	41	9.3	0.09	45	-2.58**	0.54
All teachers								
Average student standardized ELA pre-test score	-0.51	0.50	129	-0.47	0.47	128	0.52	0.07
Average student standardized math pre-test score	-0.52	0.47	129	-0.46	0.52	128	1.00	0.12
Percent students male	52.7	0.11	129	51.1	0.08	128	-1.37	0.17
Student race and ethnicity								
Percent white non-Hispanic	28.8	0.33	129	39.0	0.35	128	2.39**	0.30
Percent black non-Hispanic	47.4	0.35	129	35.9	0.34	128	-2.67***	0.33
Percent Hispanic	16.8	0.19	129	17.3	0.18	128	0.22	0.03
Percent other race	6.9	0.07	129	7.8	0.10	128	0.81	0.10
Percent students eligible for free or reduced-price lunch	76.2	0.26	115	78.0	0.20	118	0.61	0.08
Percent students English-language learners	6.5	0.10	129	7.9	0.15	128	0.83	0.10
Percent students special education	12.8	0.12	124	11.7	0.11	117	-0.74	0.10

Source: These data were provided by the district for the 2014–2015 school year for Cohort 2 and the 2015–2016 school year for Cohort 3. The sample sizes reflected in this table are for teachers who provided observation and survey data, although they may not necessarily be included in the analyses. For student information for teachers included in the analyses, see Appendix Tables E.10 and E.11.

***Indicates there are statistically significant differences between the groups at the 1 percent level, two-tailed test.

ELA = English language arts; SD = standard deviation.

2. Response rates and attrition for the one-year study

The response rate and attrition rates help provide insight into whether the observation and survey data represent the sample of teachers in the study. Having a high response rate helps determine

the study findings' potential to meaningfully represent the sample of teachers. Having low overall attrition and low differential attrition (which is the difference in the rate of attrition between the treatment and control teachers) help give confidence that the treatment and control teachers can be compared because an acceptable rate of teachers randomly assigned to treatment and control groups are represented in the analyses.

Response rates. Among novice teachers there was an overall response rate of 91 percent (88 percent control and 96 percent treatment) for novice teachers for the observations and an overall response rate of 87 percent (80 percent control and 96 percent treatment) on the survey (see Table V.6). Among all teachers there was an overall response rate of 61 percent (60 percent control and 62 percent treatment) for the observations and an overall response rate of 77 percent (69 percent control and 88 percent treatment) on the survey. The reason for the slightly lower response rates among all teachers compared to the novice teachers is because the focus of the study was on the impacts of novice teachers and additional follow-up efforts were made with novice teachers.

Table V.6. Year 1 response rates for the classroom observations and teacher survey

	Novice teachers			All teachers		
	Control (n = 65)	Treatment (n = 48)	Overall (n = 113)	Control (n = 199)	Treatment (n = 162)	Overall (n = 361)
Observation	88%	96%	91%	60%	62%	61%
Survey	80%	96%	87%	69%	88%	77%

Attrition. In addition to considering response rates, it is important to examine the rates of attrition[16] because high levels and uneven rates of attrition between the control and treatment groups can lead to the inability to compare control and treatment teachers. As described earlier, teachers were only included in our analyses if they had data on every item for a given construct, so the attrition rates and the response rates do not perfectly align. Additionally, because constructs were created using either observation or survey data, there were different rates of attrition for analyses using the observation data and analyses using the survey data. The attrition details are presented below, first for novice and all teachers included in the analyses using observation data and then for novice and all teachers included in the analyses using survey data.

Attrition in classroom observations. For all teachers, there was 37.6 percent overall attrition and 0.5 percent differential attrition. Because there was a balanced level of responses across control and treatment groups relative to the overall response rate, the results for the observation outcomes among the full sample of teachers can be interpreted as being part of a well-executed

[16] Attrition is when the teachers who were part of the schools at the time the schools were randomly assigned to treatment or control groups are not included in the impact analyses. Attrition can occur for various reasons (such as if a teacher leaves the school or does not respond to a survey).

RCT and statistically significant impacts can be attributed to TPP (What Works Clearinghouse 2017).

For novice teachers, the attrition rates indicated that disproportionately more treatment teachers contributed observation data than control teachers. The imbalance between the attrition rates of the control and treatment groups relative to the overall response rate means that the ability to compare control and treatment teachers may be compromised because a greater proportion of treatment teachers provided data compared to control teachers; therefore, the classroom observation data for the control teachers may not adequately represent the total sample of teachers who were in the control schools. To know whether we can attribute findings to TPP, we examined whether the treatment and control teachers are similar on several characteristics at baseline. As will be further described in Section C, only teachers who had a baseline measure of the outcome of interest were included in the analyses. After excluding teachers who did not meet this requirement, there was 42 percent overall attrition and 12.3 percent differential attrition among novice teachers.

Attrition in survey data. Similar to the attrition issue among novice teachers' classroom observation data, the response rates for the survey indicated that disproportionately more treatment teachers contributed survey data than control teachers.[17] We examined whether the treatment and control teachers are similar on several characteristics at baseline and only included teachers who had a baseline measure of the outcome of interest in the analyses. After excluding teachers who did not meet this requirement, there was 50 percent overall attrition and 13.7 percent differential attrition among novice teachers. For all teachers, there was 65 percent overall attrition and 16.8 percent differential attrition.

3. Sample of teachers in two-year study and description of school and students

Teachers included in the two-year analyses of TPP were part of Cohort 3 and were from 21 schools (11 control and 10 treatment). There were 55 ELA teachers (21 control and 34 treatment) who were eligible to be part of the study. Of these teachers, 23 were novice teachers (11 control and 12 treatment).

Characteristics of students taught by teachers who contributed observation and/or survey data for the two-year study were similar to those of students taught by teachers in the first year analyses. Approximately 30 percent of the students were white non-Hispanic, slightly less than 40 percent were black non-Hispanic, and less than 10 percent were from another race (see Table V.7). About a fifth of the students were of Hispanic ethnicity. Approximately half of the students were male. Over 65 percent of the students were eligible for free or reduced-price lunch. Less than ten percent of students were classified as English-language learners and less than 15 percent were

[17] The most commonly occurring attrition rates are reported for the survey data; attrition rates varied slightly across outcomes because there may be some teachers who did not provide data on one or two outcomes and, thus, were not included in the analytic sample for that outcome. The analytic samples can be found in the full results tables, Tables E.14 through E.17 in Appendix E.

classified as having special education needs. Descriptions of the characteristics of students of teachers who were included in our analyses are presented in Appendix Table E.12.

Table V.7. Comparison of characteristics of teachers' students at baseline for the two-year impact analysis

	Treatment			Control			t-statistic	Effect size
	Mean	SD	N	Mean	SD	N		
Novice teachers								
Average student standardized ELA pre-test score	-0.68	0.43	12	-1.05	0.51	7	-1.69	0.77
Average student standardized math pre-test score	-0.28	0.36	12	-0.43	0.35	7	-0.90	0.43
Percent students male	51.1	0.07	12	49.3	0.03	7	-0.66	0.32
Student race and ethnicity								
Percent white non-Hispanic	35.9	0.34	12	38.2	0.23	7	0.16	0.08
Percent black non-Hispanic	27.0	0.34	12	26.7	0.15	7	-0.02	0.01
Percent Hispanic	30.9	0.34	12	27.5	0.15	7	-0.25	0.12
Percent other race	6.20	0.05	12	7.60	0.02	7	0.75	0.36
Percent students eligible for free or reduced-price lunch	65.9	0.29	10	54.3	0.12	6	-0.93	0.48
Percent students English-language learners	8.20	0.12	12	9.0	0.06	7	0.17	0.08
Percent students special education	10.5	0.05	12	18.1	0.15	7	1.64	0.75
All teachers								
Average student standardized ELA pre-test score	-0.30	0.40	32	-0.40	0.39	17	-0.85	0.26
Average student standardized math pre-test score	-0.65	0.42	32	-0.77	0.55	17	-0.91	0.27
Percent students male	51.3	0.08	32	50.7	0.04	17	-0.27	0.08
Student race and ethnicity								
Percent white non-Hispanic	33.2	0.31	32	32.9	0.30	17	-0.03	0.01
Percent black non-Hispanic	39.4	0.37	32	44.3	0.37	17	0.45	0.13
Percent Hispanic	21.5	0.25	32	17.6	0.16	17	-0.57	0.17
Percent other race	5.9	0.05	32	5.2	0.04	17	-0.53	0.16
Percent students eligible for free or reduced-price lunch	65.4	0.35	25	69.5	0.24	13	0.38	0.13
Percent students English-language learners	6.3	0.08	32	5.1	0.06	17	-0.54	0.16
Percent students special education	13.0	0.17	32	12.0	0.12	17	-0.23	0.07

Source: The district provided these data for the 2016–2017 year for Cohort 3. The sample sizes are for teachers who provided observation and survey data, although they may not necessarily be included in the analyses. For student information for teachers included in the analyses, see Appendix Table E.12.

***Indicates there are statistically significant differences between the groups at the 1 percent level, two-tailed test.

ELA = English language arts.

We assessed baseline equivalence between treatment and control groups for novice and all teachers and interpreted statistically significant results or effect sizes greater than 0.25 as

indicating differences between the two groups. There were differences between the two groups on students' standardized ELA pre-test scores, standardized math pre-test scores, proportion male, proportion of students who were a race other than white non-Hispanic or black non-Hispanic, and proportion of students eligible for free and reduced-price lunch. Novice teachers in the control group taught students with lower ELA and math pre-test scores, fewer male students, more students of a race other than white non-Hispanic or black non-Hispanic, fewer students who qualified for free or reduced-price lunch, and more special education students. Among all teachers, there were differences between the treatment and control groups on students' standardized ELA and math pre-test scores and comparison teachers taught students with lower scores on average.

4. Response rates for the two-year study

Among novice teachers there was an overall response rate of 96 percent (91 percent control and 100 percent treatment) for novice teachers for the observations and an overall response rate of 87 percent (73 percent control and 100 percent treatment) on the survey (see Table V.8). Among all teachers there was an overall response rate of 91 percent (90 percent control and 91 percent treatment) for novice teachers for the observations and an overall response rate of 87 percent (81 percent control and 91 percent treatment) on the survey.

Table V.8. Year 2 response rates for the classroom observations and teacher survey

	Novice teachers			All teachers		
	Control (n = 11)	Treatment (n = 12)	Overall (n = 23)	Control (n = 21)	Treatment (n = 34)	Overall (n = 55)
Observation	91%	100%	96%	90%	91%	91%
Survey	73%	100%	87%	81%	91%	87%

C. Analytic approach

To address the four teacher impact research questions (see Table I.1), the impacts of TPP on teacher instructional practice constructs were estimated using regression models; specifically, we regressed each ELA instructional practice outcome on the treatment indicator and controlled for the clustering of teachers within schools, student background characteristics, and prior academic achievement (see Appendix E for more information on the regression analyses). This section first focuses on the approach to analyze impacts of one year of TPP on teachers' observed instructional practice outcomes and teachers' self-reported instructional practice outcomes for novice teachers (research question T1) and all teachers (research question T2). Then we describe the approach to analyze impacts of two years of TPP on teachers' observed and self-reported instructional practice outcomes for novice teachers (research question T3) and all teachers (research question T4). Due to the number of inferences made about ELA instructional practices, we adjusted for multiple comparisons using the Benjamini-Hochberg correction. Appendix E provides more details on the approach to estimating impacts.

1. Approach to estimate the impacts of one year of TPP for novice and all teachers (research questions T1 and T2)

Analyses were run separately for instructional practice outcomes based on observation and survey data because of different levels of attrition as described previously.

Observed teacher instructional practice outcomes. We estimated impacts of TPP on all teachers' observed ELA instructional practices using regression analyses that account for the clustering of teachers within schools and student background characteristics; this provides greater precision to the analyses and reduces the potential bias of differences between the treatment and control groups.

The observed ELA instructional practice outcomes were based on data that were combined from two time points (fall and spring) to minimize the effects of occasion variance, which is when we may not observe an instructional practice that teachers do not engage in frequently. For example, if teachers only ask students to engage in prereading activities at the start of a unit when a new text is introduced and not during every lesson, study team observers may have observed teachers during a day when the teachers did not engage in this instructional practice. By combining observations from two time points, we increase our chances of observing instructional practices that occur less frequently.

If a teacher contributed classroom observation data from both fall and spring in their first year of TPP, scores from each time point were combined to create a one-year score. If a teacher contributed classroom observation data from either the fall or the spring of their first year of TPP, scores from that one time point were used in the analyses.

As described in Section B, to estimate the effects of the TPP on novice teachers' observed ELA instructional practices, we first assessed the equivalence of the two groups prior to the intervention to have confidence that any findings are attributable to TPP and not some other preexisting characteristic. In the absence of a pre-intervention measure of teachers' instructional practices prior to their engagement with TPP—the study team used the earliest available data on teachers' instructional practices, which were collected in the fall after treatment teachers had engaged with TPP for the greater part of the fall semester. While there may have been differences in teachers' instructional practices in the fall because they had already began engagement with TPP in the summer and for part of the fall semester, these fall instructional practices data were the closest proxy for a baseline measure of treatment teachers' and control teachers' instructional practices. They were used to examine differences between treatment and control teachers and controlled for the regression analyses. Controlling for the fall baseline measure in our analyses could be interpreted as a conservative approach toward assessing impacts on spring outcomes since any differences in ELA instructional practices would be from a shorter time period than the length of the one-year TPP intervention.

Teachers who did not have data both in the fall (the proxy baseline data) and the spring (the outcome data) were excluded from the analyses.

Self-reported teacher instructional practice outcomes. We estimated impacts of TPP on novice teachers' and all teachers' self-reported instructional practices using regression analyses that account for the clustering of teachers within schools and student background characteristics.

Similar to the approach to estimate impacts among novices on observed instructional practice outcomes, we assessed the equivalence of the two groups prior to the intervention to using the fall data, which were the earliest available data on teachers' instructional practices. Teachers who did not have data both in the fall (the proxy baseline data) and the spring (the outcome data) were excluded from the analyses.

2. Approach to estimate the impacts of two years of TPP for novice and all teachers (research questions T3 and T4)

Impacts of two years of TPP on Cohort 3 novice teachers' and all teachers' instructional practices were estimated using regression analyses that account for the clustering of teachers within schools and—given the small sample size—select student background characteristics: ELA and math pre-test scores, English-language learner status, and special education status. Given that this study follows a quasi-experimental design, it is necessary to test for baseline differences on an early measure of the instructional practice of interest. To do so, we used the teachers' instructional practice data from the fall of their first year (fall 2016) and used data from the 2017–2018 school year as their instructional practice outcome data. Teachers were excluded from the analysis if they did not have fall 2016 data and data from either fall 2017 or spring 2018. If teachers contributed data in both fall 2017 and spring 2018, these data were averaged and used in the analyses.

Due to smaller sample sizes in the second year of TPP, it was not possible to include all the baseline characteristics in the regression analyses. The analyses adjusted for only a subset of students' background characteristics that would affect teachers' practices: students' special education status, English-language learner status, standardized ELA pre-test score, and standardized math pre-test score—but not race, Hispanic ethnicity, free or reduced-price lunch, and proportion male. The analyses adjusted for students' special education status and standardized ELA and math pre-test scores because there were the largest differences between the treatment and control groups on these characteristics. The analyses adjusted for the proportion of English-language learners because of the explicit and structured instruction required to support students learning English (Baker et al. 2014). In addition, the analyses adjusted for teachers' instructional practices at baseline, or the fall of the prior year. For the two-year impact analyses, because of low statistical power due to small sample sizes, the effects of TPP would need to be very large for the analyses to be able to detect any significant impacts. Findings should be interpreted with caution because low power may explain the lack of statistically significant results and not the effects of TPP.

D. Findings

This section presents the results from our analyses of TPP on teachers' instructional practices after one year of TPP and after two years of TPP and how the results compared to the findings

from earlier research. In the first section, we present the findings of TPP after one year on novice and all teachers' observed instructional practice constructs and then their self-reported instructional practice constructs. The second section describes the findings of TPP after two years on novice and all teachers' instructional practice constructs. The findings based on observations and self-reports are presented in tandem because of the similar approach for the two-year analyses. This section summarizes the findings at a high level with focus on findings with statistically significant impacts; the full set of results covering all ELA instructional practice outcomes are reported in Appendix Tables E.14–E.17.

1. Impacts of one year of TPP on teachers' ELA instructional practices (research questions T1 and T2)

We found that one year of TPP had statistically significant, positive impacts on novice and all treatment teachers' having connected lessons and their overall ELA instructional practices (see Table V.9). Additionally, there was a statistically significant, positive impact of one year of TPP on several outcomes among all teachers: supporting students' higher-order thinking; supporting students' participation; engaging students in reading, writing, and/or speaking about texts; and supporting students' use of text evidence.

In the next sections, we present the impacts of one year of TPP for individual constructs; first, we show a summary of whether or not TPP had a significant impact for topics in each area; then we present the impacts for each outcome within the topics and discuss these findings alongside those from the earlier analyses.

Table V.9. Summary of impacts of one year of TPP, by topic

Main areas	Topics	Number of constructs	Were there significant impacts of TPP?	
			Novice teachers	All teachers
General instruction	Connected lessons	2	✓	✓
	Connections to world	1		
	Content knowledge	2		
	Higher-order thinking	3		✓
	Multimedia use	2		
	Prior knowledge	1		
	Student independence	1		
	Student participation	3		✓
	Student responsibility for learning	1		
Reading and writing instruction	Academic vocabulary	1		
	Multiple types and purposes of writing	2		
	Reading, writing, and/or speaking about texts	3		✓
	Use of text evidence	3		✓
	Writing conventions	2		
Classroom management and environment	Classroom climate	1		
	Classroom management	3		
Overall ELA instructional practices		2		✓

Note: ✓ indicates that, for that topic, there was one construct with a statistically significant, positive impact of TPP after adjusting for multiple comparisons.

ELA = English language arts.

a. General instruction practices

Key findings: One year of TPP showed positive, statistically significant impacts for all teachers and novice teachers having clearly stated learning purposes in their lessons. One year of TPP showed positive, statistically significant impacts for all teachers supporting students' higher order thinking, engaging students in collaborative discussion practices, and encouraging students' participation in discussions.

One year of TPP had statistically significant, positive impacts for novice and all teachers on the general instructional practice of making the learning purpose of the lesson clear, such as by stating the purpose of instruction out loud and referring back to the learning goals during the lesson.

Because of their TPP engagement, there were impacts on all teachers' practices of supporting students' higher-order thinking through activities, such as teachers supporting students' analysis of text and asking questions to encourage students to think at a higher level. This finding parallels ones from our earlier research in which a larger proportion of TPP teachers demonstrated practices that supported their students' critical thinking skills during classroom discussions compared to the control teachers (Choi et al. 2018, Appendix F). While the current analyses found no detectable impacts on constructs reflecting novice teachers' practices of supporting students' higher-order thinking skills, prior research indicated that a larger proportion of novice TPP

teachers did demonstrate some specific practices that engaged students in higher order thinking skills compared to control novice teachers (Choi et al. 2017, Choi et al. 2018, Appendix F).

In addition to TPP impacts on teachers' supporting students' higher-order thinking, findings from this report's analyses showed statistically significant, positive impacts of one year of TPP on all teachers' practices to support collaborative discussions among students and encouraging student participation in discussion (see Table V.10). This means that, compared to the control teachers, all teachers who engaged in TPP for a year used these practices more—practices that included supporting students to respond to peer comments and build on each other's thinking and asking other students to help answer questions in discussions.

In addition, among novice teachers, one year of TPP had potentially statistically significant, positive impacts of teachers having connected lessons. This finding indicates that TPP teachers had connected lessons to a greater degree than control teachers. However, we are unable to interpret the differences as fully attributable to TPP because the treatment and control teachers showed early differences in their use of instructional practices in the fall, which indicates there may be additional reasons for the finding beyond teachers' TPP engagement between the fall and spring.

There were no statistically significant impacts of one year of TPP on any other general ELA instructional practice outcomes for novice or all teachers, as shown in Appendix Tables E.14 and E.15, which report results from all the teacher instructional practice outcomes, including all insignificant findings. Although the findings from the current analyses did not demonstrate impacts of TPP on teachers developing students' content knowledge, earlier research indicated that a larger proportion of novice TPP teachers developed students' content knowledge through their lessons and supported students in taking responsibility for their own learning compared to control teachers (Choi et al. 2017).

Key findings:

- One year of TPP showed statistically significant, positive impacts for novice and all teachers' practices of **supporting students' close reading and writing that demonstrates understanding of the text**.
- One year of TPP showed statistically significant, positive impacts for novice teachers' practice of **supporting students' use of evidence from text and engagement in argument writing**.
- One year of TPP showed potentially statistically significant, positive impacts for all teachers' practice of **supporting students' use of evidence from texts.**

b. CCSS-aligned reading and writing instructional practices

There were statistically significant, positive impacts on CCSS-aligned instructional practices of all teachers supporting students' close reading and writing to demonstrate understandings of text(s). Though the current analyses did not show impacts of TPP on novice teachers' supporting close reading and writing to demonstrate understandings of text(s), findings from prior research indicated that more novice TPP teachers engaged students in conducting close readings of the text compared to control novice teachers (Choi et al. 2017, Appendix F).

One year of TPP also had potentially statistically significant, positive impacts for all teachers supporting students' use of evidence from texts. This finding indicates that TPP teachers

supported students' use of evidence from texts to a greater degree than control teachers. However, due to early differences in their use of instructional practices in the fall, we are unable to interpret the finding as fully attributable to the TPP because the early differences indicate there may be additional reasons for the finding beyond teachers' TPP engagement between the fall and spring. Earlier research indicated that a larger proportion of all TPP teachers and novice TPP teachers had students use evidence from texts by rereading the text and citing evidence from the text to support their responses compared to all control and all novice teachers, respectively (Choi et al. 2017, Choi et al. 2018, Appendix F).

There were no statistically significant impacts of one year of TPP on any other reading and writing instructional practice outcomes for novice or all teachers, as shown in Appendix Tables E.14 and E.15, which report results from all the teacher instructional practice outcomes. While there were no impacts on novice or all teachers' practices of supporting students' reading, writing, and speaking about texts based on analyses for this report, findings from earlier research are that a larger proportion of all TPP teachers, including novice TPP teachers, engaged students in these activities than control teachers (Choi et al. 2017, Choi et al. 2018, Appendix F).

c. Classroom management and environment

There were no statistically significant impacts of one year of TPP on any ELA instructional practice outcomes related to classroom management and environment for novice or all teachers, as shown in Appendix Tables E.14 and E.15, which report results from all the teacher instructional practice outcomes, including all nonsignificant findings.

d. Impacts of overall ELA instructional practices

Key findings: A single year of TPP showed positive, statistically significant impacts on all teachers' **overall ELA instructional practices.**

A single year of TPP had positive, statistically significant impacts on all teachers' overall ELA instructional practices. This means that all teachers who engaged in TPP for a year used more of the overall ELA instructional practices than teachers who continued with their district-provided PD and curriculum. The overall ELA instructional practices outcome included the general instruction, reading and writing instruction, and classroom management and environment areas. The practices represented by these areas included CCSS-aligned practices, like supporting students' academic vocabulary development, higher-order thinking, use of evidence from texts, and content knowledge development. This practices in these areas also included foundational instructional practices like supporting collaborative student discussion and classroom management practices.

One year of TPP also had potentially statistically significant, positive impacts for novice teachers' overall ELA instructional practices. Due to early differences in the instructional practices between TPP and control novice teachers, it is unclear whether the finding is fully attributable to TPP because there may be other reasons for the differences beyond teachers' TPP engagement between the fall and spring. However, the statistically significant finding for the full sample of teachers, which includes the novice teachers, demonstrates the positive impacts of TPP on teachers' instructional practices after one year.

Table V.10. Highlighted one-year impacts on select English language arts teachers' instructional practices

Topic	Outcome	Data source		Estimated treatment coefficient	Standard error	Effect size	Analytic sample size	
		Observation	Survey				Treatment	Control
Novice teachers								
General instruction								
Connected lessons	Connected lessons	x		0.22*	0.07	1.27	31	34
	Stated learning purpose[a]	x		0.33*	0.10	2.92	31	34
Overall ELA instructional practices								
Overall ELA instructional practices[a]		x		1.78*	0.47	1.42	31	34
			x	-0.17	0.40	-0.18	24	25
All teachers								
General instruction								
Connected lessons	Connected lessons	x		0.07	0.03	0.69	75	89
	Stated learning purpose	x		0.32*	0.05	3.53	75	89
Higher order thinking	Assessment of higher order thinking		x	0.02	0.04	0.21	72	55
	Higher order thinking and reading		x	0.01	0.04	-0.15	71	55
	Observed higher order thinking	x		0.13*	0.03	1.06	75	89
Student participation	Collaborative discussion practices	x		0.12*	0.03	1.24	75	89
	Student engagement in class	x		-0.02	0.03	-0.39	75	89
	Student participation in discussion	x		0.08*	0.03	0.79	75	89
Reading and writing instruction								
Reading, writing, and/or speaking about texts	Close reading and writing that demonstrates understanding of texts	x		0.10*	0.04	1.48	75	89
	Reading, writing, and/or speaking about texts in past two weeks		x	-0.02	0.04	-0.16	72	54
	Reading, writing, and/or speaking about texts in typical week		x	0.01	0.04	0.20	72	55
Use of text evidence	Assessment of text evidence use		x	0.00	0.04	-0.19	72	55
	Self-reported text evidence use[a]		x	0.11*	0.04	0.78	71	55
	Text evidence use and argument writing	x		0.09	0.04	1.38	75	89
Overall ELA instructional practices								
Overall ELA instructional practices		x		1.19*	0.24	1.30	75	89
			x	-0.08	0.31	-0.05	38	38

Table V.10 (*continued*)

Source: For all teachers' observation outcomes, impacts were estimated on teachers' practices in 2015–2016 for Cohort 2 schools and 2016–2017 for Cohort 3 schools. For novice teachers' observation outcomes and novice and all teachers' survey outcomes, impacts were estimated on teachers' practices in spring 2016 for Cohort 2 schools, accounting for fall 2015 as the baseline measure, and spring 2017 for Cohort 3 schools, accounting for fall 2016 as the baseline measure.

Note: This table provides results for topic areas only if there were one or more statistically significant impacts on a teacher instructional practice outcome. Appendix Tables E.14 and E.15 show the full results.

Effect size is the adjusted difference between treatment and control group average teacher instructional practice scores divided by the unadjusted pooled standard deviation, with an adjustment for small sample sizes per What Works Clearinghouse guidance.

For observation outcomes, the following covariates were included in the regression models: standardized ELA pre-test score, standardized math pre-test score, male, race (white non-Hispanic, black non-Hispanic, and other), Hispanic ethnicity, free or reduced-price lunch eligibility, English-language learner status, and special education status.

For survey outcomes, the following covariates were included in the regression models: teacher instructional practice of interest at baseline, standardized ELA pre-test score, standardized math pre-test score, male, race (white non-Hispanic, black non-Hispanic, or other), Hispanic ethnicity, free or reduced-price lunch eligibility, English-language learner status, and special education status.

*Significant differences at the 5 percent level between control and treatment teachers after correcting for multiple comparisons.

[a]Result should be interpreted with caution since there were differences between the groups at baseline on this instructional practice.

ELA = English language arts.

There were no statistically significant impacts of one year of TPP on the self-reported overall ELA instructional practice outcome for novice or all teachers, as shown in Appendix Tables E.14 and E.15, which report results from all the teacher instructional practice outcomes, including all insignificant findings.

2. Impacts of two years of TPP on teachers' instructional practices (research questions T3 and T4)

There were no statistically significant impacts of two years of TPP on novice and all teachers' ELA instructional practices, as can be seen in Appendix Tables E.16 and E.17. The earlier research of the effects of TPP among all teachers after one and two years indicated that significantly more TPP teachers encouraged students' higher-order thinking, asked students to use evidence from texts, and engaged more often in reading, writing, and speaking about texts compared to teachers who did not engage in TPP (Choi et al. 2018, Appendix F).

E. Considerations when interpreting the findings

Through the analyses conducted for this report, we examined the one-year and two-year impacts of TPP on teachers' ELA instructional practice outcomes. When interpreting the findings from these analyses, there are some key considerations to keep in mind.

With respect to the findings from the one-year impacts of TPP on the observed ELA instructional practices among novice teachers and the self-reported ELA instructional practices among all teachers and novice teachers, we assessed the baseline equivalence of the treatment and control teachers using teachers' fall ELA instructional practice scores. Because the fall data were collected after the treatment teachers had engaged in TPP in the summer and for part of the fall semester, any preexisting differences between the treatment and control teachers may be at least partly due to the early effects of TPP. However, because of differences in attrition between the treatment and control groups, it is also possible that any preexisting differences may be at least partly due to differences in the groups' background characteristics. We are unable to determine whether the differences in earlier instructional practices among the treatment and control teachers are attributable to TPP, differences in background characteristics, or other factors.

Additionally, for the two-year impacts of TPP on observed and self-reported ELA instructional practices, teachers who did not have data on the ELA instructional practices both in the fall 2016 (the baseline data) and in the 2017–2018 school year (the outcome data) were excluded from the analyses. This meant that of the smaller subset of districts, schools, and teachers who participated in the two-year impact study, an even smaller subset of teachers were included in the analyses. This resulted in an analysis of 15 to 18 novice teachers and between 21 to 35 teachers in the full sample. With such small sample sizes, the analysis is underpowered to detect differences among groups while controlling for background characteristics at baseline. For this reason, readers should exercise caution when interpreting the results as the lack of significant results may be due to the small sample sizes and not TPP.

VI. IMPACTS OF TPP ON STUDENT ACHIEVEMENT

The ultimate goal of supporting teachers to deliver high-quality instruction is to improve student achievement. TPP's supports for English language arts (ELA) teachers aim to enhance student learning in ELA and literacy. To determine the impact of TPP on the ELA achievement of the students of study teachers, we analyzed two measures of student performance: test scores on state ELA achievement tests at the end of the first and second years of schools' participation in the study, and scores on a literacy task developed and administered in spring 2017 of the study. The analysis and findings on these two outcome measures are presented in turn.

A. Impacts on student ELA test scores

In this section, we first discuss the test score measures used. We next describe the analytic approach and the sample. Finally, we present the findings, including considerations for interpretation.

1. ELA test score outcome measures

In this study, ELA achievement was assessed using data from the statewide ELA exams administered by the participating school districts. Although the specific tests differ across states,[18] the ELA exams in general are designed to capture information about students' reading and some information about students' writing achievement. In these tests, students often answer multiple-choice questions based on short passages they read, and write responses to open-ended questions based on stories, articles, or poems they listen to or read. For example, the English language arts/Literacy portion of the PARCC assessment includes narrative, research, and literary analysis tasks.

In order to make ELA test scores comparable across the states, we constructed standardized test scores of study students relative to the average student nationally. In particular, the ELA scores were converted to *z*-scores by subtracting the mean and dividing by the standard deviation of a national norming population within each grade and year, published by test publishers.[19]

[18] Tests include ACT Aspire, End of Grade Testing (EOG), Massachusetts Comprehensive Assessment System, Milestones EOG, Missouri Assessment Program (MAP), New York State assessment, Partnership for Assessment of Readiness for College and Careers (PARCC), Scantron, Smarter Balanced, and Student Assessment of Growth and Excellence.

[19] For schools in two districts and one school in another district, we could not obtain published means and standard deviations and had to use control sample means and standard deviations within state and grade. As a rule, we use national-level means and standard deviations for states that took national-level exams, such as PARCC, but we use state level means and standard deviations for exams that are only offered in one state, such as the New York State assessment.

2. Analytic approaches

We estimated a number of different impacts to address research questions about the effects of TPP on students, shown in Table I.1. The analyses included subsets of 4th- through 8th-grade students of novice and all teachers in the first and second year of TPP implementation. Below, we describe each of these types of impacts and introduce terminology, then describe our impact estimation methods.

a. One-year impacts

First, we estimated impacts of TPP after one year of implementation to learn whether it led to increases in student achievement by the end of that year. These one-year impacts involved teachers who had one year of exposure to TPP and the students who were in their classroom that year. We examined the ELA achievement of the students of novice study teachers and all study teachers using the full sample of schools in the study. We refer to these as the one-year novice impact (addressing research question S1) and one-year impact (addressing research question S2), respectively. In these analyses, we compare students in the classrooms of study teachers in treatment schools at the end of the school's first year in the study with students in the classrooms of control teachers at the same point in time. We refer to the samples of students used to estimate the one-year novice impact and the one-year impact as the one-year novice impact and one-year impact samples.

b. Two-year impacts

Next, we estimated several impacts after two years to learn about how the effects of different types of engagement with TPP evolve over time. These analyses have in common a comparison of outcomes among students in the classrooms of ELA teachers in schools' second year in the study.

- A first type of two-year impact is the average impact among all study schools after two years, referred to as the overall two-year impact (addressing research question S4). The overall two-year impact assesses the influence of at least one year of engagement with TPP by comparing students of study teachers in schools that received either one or two years of TPP to students of study teachers in schools that did not receive TPP.

- A second type of two-year impact is the impact of two years of implementation of TPP among the students of novice study teachers and all study teachers, referred to as the two-year novice impact (addressing research question S5) and two-year impact (addressing research question S6), respectively. These analyses compare students of study teachers in the selected schools that received two years of TPP at the end of that second year to analogous students of study teachers in schools that did not receive TPP.

- A third type of two-year impact, referred to as the two-year schoolwide impact (addressing research question S7), reflects the effect of an entire study school engaging with TPP. The analysis compares students of all ELA teachers in the selected schools that received two

years of TPP—including teachers who did not receive TPP in addition to study teachers who received two years of TPP— at the end of that second year to analogous students of all teachers in schools that did not receive TPP.

All three types of two-year impacts reflect the impact on student achievement over the course of one year (a school's second year in the study).

c. Follow-up impacts

Finally, we estimated the impact of a single year of implementation of TPP on students in the following year to learn whether a single year of engagement has impacts that persist over time in the absence of continuing TPP support. The analyses compare students of study teachers in the selected schools that received one year of TPP services during the previous study year to students of study teachers in schools that did not receive TPP, separately for novice teachers and all teachers. We refer to these impacts as the one-year follow-up novice impact (addressing research question S8) and one-year follow-up impact (addressing research question S9), respectively. Because this group of schools is also included in the sample used to estimate the overall two-year impact, together with schools that received two years of TPP, these contrasts will help us to understand to what extent two-year effects were driven by schools that received two years of TPP versus lingering effects among the schools that received only one year of TPP.

d. Impact estimation methods

To estimate each impact of TPP on student ELA achievement, we used a benchmark regression model that analyzed standardized test scores as a function of a treatment indicator as well as baseline characteristics of schools, teachers, and students as covariates in regression analyses. In particular, baseline student characteristics included standardized ELA and math test scores from the prior year (pre-test scores),[20] gender, race and Hispanic ethnicity, free or reduced-price lunch (FRPL) eligibility, English-language learner (ELL) status, and special education (SPED) status. Teacher characteristics included an indicator of novice status. School characteristics included charter status, number of students, proportion of students eligible for free and reduced-price lunch, and number of full time teacher equivalents. The model also included district-by-grade and year indicators. Missing values of covariates (including pre-test scores) were set to a single value (0 for binary variables and the sample mean for continuous variables), and indicators for missing values of covariates were included as controls in the regression.

Regression analyses used nonresponse weights since some students did not have test score outcome data. These weights were based on the inverse of the predicted nonresponse based on students' demographic characteristics (gender, race and ethnicity, FRPL, SPED, and ELL) and an indicator of school grade configuration (elementary or middle) so that students whose

[20] Note that for the two-year estimates (addressing research questions S4–S7), using the prior year baseline provided conservative estimates of TTP impacts for those students who were exposed to two years of a TPP teacher. In other words, the prior-year baseline isolated impacts of just the most recent year since it controlled for impacts through the end of the previous year.

characteristics were more similar to nonresponders were given more weight in analysis. Standard errors accounted for the clustering of students within schools.

3. Description of student samples

In this section, we describe the students in the samples used to estimate two key impacts described above: the one-year impact and the two-year impact. We describe the sample sizes and characteristics of these students, and present information about attrition and baseline equivalence as relevant. Appendix tables present this information for samples used in all the student impact analyses described previously in Section 2. Appendix Tables G.3–G.4 present information about sample sizes and demographics, while Appendix Tables G.1, G.2, G.5, and G.6 present information about attrition and baseline equivalence for all student analyses conducted for the study.

The ELA achievement analyses pool 66 schools in all three cohorts that provided useable test score data (representing 61 out of 70 participating study schools).[21,22] Within these schools, we examine the sample of students who were enrolled both in October of a given school year and at the time tests were administered in the spring.[23] We imposed several restrictions on the student sample. We omitted those students who left or joined during the school year. Students whose teachers had five or fewer students were dropped from analysis because they are likely to be specialized classes that differ from the average ELA class; students who skipped or repeated grades were dropped because their pre-tests or post-tests differ from those of the other students (for example, a sixth-grade student who skipped grade 5 would have a grade 4 pre-test and a grade 6 post-test, unlike other sixth-grade students).

a. One-year sample

Below we describe the one-year impact sample, the largest analytic sample used in the study. To capture the one-year intervention impact for students across all schools involved, we compare students in the classrooms of study teachers in treatment schools at the end of the school's first year in the study with students in the classrooms of control teachers at the same point in time. This sample included students in the classrooms of study teachers at the end of schools' first year of TPP implementation. These students were from 66 schools across three cohorts, resulting in a final sample of 12,859 students for one-year impact analyses. 34 schools were in the treatment group and contained 6,150 students; 32 schools were in the control group and contained 6,709 students. These students were taught by 134 treatment teachers and 134 control teachers, including 87 novice teachers (40 treatment and 47 control). In terms of student demographics, 49 percent were female, 38 percent were white non-Hispanic, 35 percent were black non-Hispanic,

[21] These are counts of the schools as they were structured at random assignment.

[22] Useable data means both years of student data were provided in time to be included in the analyses. Two study districts did not contribute data to the main or sensitivity analyses, as described in Chapter III.

[23] All schools in one district and one school in another district only provided end of year rosters and so were not included in the main impact analysis but were included in a sensitivity analysis.

8 percent reported another race, 20 percent were Hispanic and 60 percent were eligible for free or reduced-price lunch (Appendix Table G.3).

We examined the sample loss after random assignment (attrition) from the one-year impact samples (Appendix Tables G.1 and G.2). The purpose of random assignment is to produce a treatment group and a control group that are equivalent on all characteristics at the start of the study. If the two groups are indeed equivalent at the outset, and if any attrition from the sample over the course of the study is low overall and balanced across groups, one can be confident that any differences in outcomes between the two groups found later are due to the intervention. For the one-year impact sample (addressing research question S2), the overall school attrition rate was 23 percent, and the differential attrition between the two experimental conditions was 0.6 percentage points. The overall student attrition rate was 17 percent, and the differential attrition rate was 2 percentage points. Attrition was low for the one-year novice impact sample (addressing research question S1) as well: overall and differential school attrition were 37 percent and 7 percent, while overall and differential student attrition were 12 percent and 1.3 percent, respectively. These combinations of overall and differential attrition for the one-year impact samples fall within the acceptable "optimistic" What Works Clearinghouse threshold for low attrition.

b. Two-year sample

The overall two-year impact sample was the largest analytic sample used in the study for two-year analyses. To capture the overall two-year intervention impact for students across all schools involved, we compared a treatment sample of students taught by treatment teachers that received one or two years of TPP services and were in the treatment schools for two years to a control sample of students taught by control teachers who were in the control schools for two years, which makes the group of control teachers included in the analysis consistent with the treatment teachers. These students were from 66 schools across three cohorts, resulting in a final sample of 10,159 students for the overall two-year impact analysis. 34 schools were in the treatment group and contained 5,168 students; 32 schools were in the control group and contained 4,991 students. These students were taught by 104 treatment teachers and 95 control teachers. In terms of student demographics, 50 percent were female, 34 percent were white non-Hispanic, 33 percent were black non-Hispanic, 9 percent reported another race, 24 percent were Hispanic and 61 percent were eligible for free or reduced-price lunch (Appendix Table G.4).

For the two-year impact analyses equivalence of the intervention and comparison groups on observable characteristics at baseline must be established for the analytic samples. In Appendix Table G.5, we report the effect sizes (standardized mean differences) in students' baseline ELA test scores. According to What Works Clearinghouse guidance, effect sizes greater than 0.25 indicate that differences between the treatment and control groups on baseline ELA measures cannot be accounted for with a statistical adjustment, and therefore the groups cannot be considered equivalent. Effect sizes of larger than 0.05 and smaller than 0.25 indicate a difference that requires statistical adjustment. As Appendix Table G.5 shows, all student ELA baseline differences are within the adjustment range (less than 0.25 standard deviations and greater than 0.05 standard deviations) and, therefore, ELA pre-test scores are included as covariates in all

impact models. In addition to the baseline ELA measures that are required for satisfying the baseline equivalence requirement, other sample characteristics such as grade level may be associated with the outcome. In Appendix Table G.6, we report counts and proportions of students enrolled in each grade (4–8) by experimental condition for study analytic samples. Large baseline differences (greater than 0.25 standard deviations) in grade level among students of novice teachers could affect analyses of impacts on the students of novice teachers and interpreted as evidence that the treatment and control groups of students are not sufficiently comparable. Therefore, student findings for novice teachers addressing research questions S1, S5, and S8 should be interpreted with caution.

4. Findings

In this section, we discuss the impacts of TPP on students' ELA achievement. To assess whether treatment students outperformed control students, Table VI.1 provides results from a benchmark analytic model for the eight contrasts defined by the research questions shown in Table I.1. The effect size indicates the magnitude of the estimated impact, calculated as a proportion of the pooled standard deviation of the outcome measure for the treatment and control group.

Table VI.1. Student achievement impacts of TPP

Impact and research question addressed	Estimated treatment coefficient		Standard error	p-value	Effect size	Analytic sample size	
						Control	Treatment
One-year impacts in all schools							
(1) One-year novice impact (S1)	-0.003		(0.02)	0.89	**0.00**	2,653	2,162
(2) One-year impact (S2)	0.015		(0.01)	0.25	**0.01**	6,709	6,150
Overall two-year impacts in all schools							
(3) Overall two-year impact (S4)	0.084	***	(0.02)	< 0.01	**0.08**	4,991	5,168
Two-year impacts in schools that participated in the two-year study							
(4) Two-year novice impact (S5)	0.047		(0.05)	0.37	**0.05**	738	1,063
(5) Two-year impact (S6)	0.092	***	(0.03)	< 0.01	**0.10**	1,145	2,631
(6) Two-year schoolwide impact (S7)	0.017		(0.03)	0.32	**0.02**	3,374	3,486
Follow-up impacts in schools that did NOT participate in the two-year study							
(7) One-year follow-up novice impact (S8)	0.017		(0.13)	0.89	**0.02**	1,161	555
(8) One-year follow-up impact (S9)	0.055	*	(0.03)	0.07	**0.06**	3,846	2,537

Note: Seventeen districts contributed data to the student achievement impact analysis. The one-year impact analyses exclude one district, while the two-year analyses exclude two districts. Effect size is the adjusted difference between treatment and control group average standardized test scores divided by the student level unadjusted pooled standard deviation of post-test scores.

*Statistically significant at the 10 percent level, two-tailed test.
**Statistically significant at the 5 percent level, two-tailed test.
***Statistically significant at the 1 percent level, two-tailed test.

a. One-year impact findings

Key findings:
- TPP did not have a statistically significant impact on student ELA achievement after one year of implementation.

One-year novice impacts (row 1, addressing research question S1) and one-year impacts (row 2, addressing research question S2) indicate that students of teachers in the treatment schools that received one year of the TPP program did not perform better than students in the control schools, regardless of teachers' professional experience. Effect sizes of 0 and 0.01 are not statistically significant. Similarly, students of novice teachers that received one year of TPP services during the *previous* study year performed no better than students of control teachers. The one-year follow-up novice impact, shown in row 7 (addressing research question S8), indicates that the effect size of 0.02 is not statistically significant. However, the one-year follow-up impact, which encompassed all teachers, has an effect size of 0.06 that approaches statistical significance at the 5 percent level. The fact that this effect is detected after TPP was discontinued in those schools suggests the persistence and gradual maturation of the teacher PD intervention.

b. Two-year impact findings

Key findings:
- TPP had a positive and statistically significant impact on student ELA achievement after two years among all study schools that received either one or two years of TPP.
- TPP had a positive and statistically significant impact on student ELA achievement after two years among schools that implemented TPP for two years.

We examined outcomes in two-year impact analyses in order to assess the impact of TPP after the longest time period possible in the study. Even though there were no impacts of TPP on student achievement after one year, impacts emerged after two years. We found a positive and statistically significant overall two-year impact among all study schools that received either one or two years of TPP, with an effect size of 0.08 (row 3, addressing research question S4). This effect size indicates that 53 percent of treatment group students can be expected to perform higher on ELA than the average comparison group student as a result of the intervention. It corresponds to roughly 1.1 months of growth in student achievement (assuming 0.7 standard deviations of improvement over a 10-month school year). Note that, mathematically, this overall two-year impact is roughly an average of the impacts among schools that received two years of TPP and those schools that received one year of TPP services during the previous study year because it combines both of these subgroups of schools.

We also found a positive and statistically significant two-year impact among schools that implemented TPP for two years, with an effect size of 0.10 (row 5, addressing research question S6). This effect size indicates that 54 percent of the intervention group can be expected to perform higher on ELA than the average comparison group student as a result of the intervention. It corresponds to roughly 1.4 months of growth in student achievement. These findings appear to be driven by TPP and not by a more proficient sample of teachers in these schools; we found no impacts after one year among the teachers that received two years of TPP.

The two-year novice impact (row 4, addressing research question S5) and the two-year schoolwide impact of TPP (row 6, addressing research question S7) were both positive but not statistically significant, indicating that TPP did not have a statistically significant impact on the students of novice teachers or the students of all teachers (whether included in the study or not) during schools' second year of TPP.

Because we tested multiple hypotheses about student achievement, we adjusted for multiple comparisons across all analyses in Table VI.1 using the Benjamini-Hochberg correction. The adjustment did not change the statistical significance of the two main findings. Both the overall two-year impact and two-year impact findings (addressing research questions S4 and S6) remain statistically significant at the 5 percent level after this adjustment.

Key findings:
- TPP did not have a statistically significant impact on the ELA achievement of students of novice teachers after one or two years of implementation.

5. Sensitivity analyses

In this section, we discuss the findings from analyses investigating the sensitivity of our results to different analytic model specifications. To examine robustness of the findings, several approaches to sensitivity were taken. We first focus on an alternative method of addressing missing data: complete case analysis. Next, we examine analyses that include student joiners who may have had more limited exposure to the intervention. Finally, we describe a number of other sensitivity analyses in Appendix G. Overall, the findings are robust to various specifications and do not change qualitatively.

a. Complete case analysis

In Table VI.2, we first present findings for complete case analysis that differs from the benchmark analysis in its treatment of missing data. The benchmark analysis sets all missing values for a baseline measure to a single value and includes an indicator variable for records missing data on the measure in the impact estimation model. The complete case analysis (commonly known as "listwise deletion") excludes observations with any missing outcome and baseline data from the analysis. Compared to the benchmark one-year impact analysis (addressing research question S2) in Table VI.1, the complete case analysis excluded 3,202 students with missing outcome data. This did not cause high differential attrition rates by program condition: the analysis kept 73 percent of the total sample of 6,150 treatment students and 77 percent of the 6,709 controls for the sensitivity analyses, for differential attrition of 4 percent. As shown in Table VI.2, the impact estimates for the complete case analyses were somewhat larger in magnitude than those for the benchmark analyses but were similar in statistical significance. The overall two-year impact (addressing research question S4) remains significant and the effect size increased slightly from 0.08 to 0.09 standard deviations. The two-year impact (addressing research question S6) remains significant and the effect size increased from 0.10 to 0.14 standard deviations, the largest increase in effect sizes among the complete case analyses. The one-year follow-up impact finding (addressing research question S9) decreases and no longer approaches the 5 percent level of significance.

Table VI.2. Complete case sensitivity analysis: student achievement impacts of TPP

Impact and research question addressed	Estimated treatment coefficient		Standard error	p-value	Effect size	Analytic sample size	
						Control	Treatment
One-year impacts in all schools							
(1) One-year novice impact (S1)	–0.008		(0.02)	0.74	**–0.01**	2,239	1,693
(2) One-year impact (S2)	0.004		(0.01)	0.75	**0.00**	5,151	4,506
Overall two-year impacts in all schools							
(3) Overall two-year impact (S4)	0.090	***	(0.02)	< 0.01	**0.09**	3,963	3,635
Two-year impacts in schools that participated in the two-year study							
(4) Two-year novice impact (S5)	0.056		(0.05)	0.24	**0.06**	609	787
(5) Two-year impact (S6)	0.130	***	(0.03)	< 0.01	**0.14**	821	1,787
(6) Two-year schoolwide impact (S7)	0.030		(0.02)	0.10	**0.03**	2,485	2,473
Follow-up impacts in schools that did NOT participate in the two-year study							
(7) One-year follow-up novice impact (S8)	–0.032		(0.13)	0.81	**–0.03**	887	413
(8) One-year follow-up impact (S9)	0.038		(0.03)	0.22	**0.04**	3,142	1,848

Note: Seventeen districts contributed data to the student achievement impact analysis. The one-year impact analyses exclude one district, while the two-year analyses exclude two districts. Effect size is the adjusted difference between treatment and control group average standardized test scores divided by the unadjusted pooled standard deviation.

*Statistically significant at the 10 percent level, two-tailed test.

**Statistically significant at the 5 percent level, two-tailed test.

***Statistically significant at the 1 percent level, two-tailed test.

b. Inclusion of student joiners

In schools, students transfer out and new students enter into schools continually throughout the year. Student mobility, a phenomenon over which schools exercise little control, is especially widespread in schools serving students from low-income families. Selective flow into and out of the study sample can bias results. In Table VI.3 we present a sensitivity analysis that differs from the benchmark analyses in its inclusion of "joiners" in the analytic sample. That is, the analytic sample of students was identified at the time outcomes were measured in the spring rather than in October, and includes students who joined schools during the school year. This analysis expands the student sample used in the benchmark analysis to include those who may have been exposed to the treatment for only part of the school year. For one-year impacts (addressing research question S2), for example, the combined sample included the benchmark sample of 12,859 students (Table VI.1) and 1,973 additional students who had moved into the treatment and control schools after the baseline assessments.

Table VI.3. Inclusion of student joiners' sensitivity analysis: school-level achievement impacts of TPP

Impact and research question addressed	Estimated treatment coefficient		Standard error	p-value	Effect size	Analytic sample size	
						Control	Treatment
One-year impacts in all schools							
(1) One-year novice impact (S1)	0.019		(0.02)	0.42	**0.02**	2,822	2,453
(2) One-year impact (S2)	0.002		(0.01)	0.85	**0.00**	7,866	6,966
Overall two-year impacts in all schools							
(3) Overall two-year impact (S4)	0.087	***	(0.02)	< 0.01	**0.09**	5,116	5,397
Two-year impacts in schools that participated in the two-year study							
(4) Two-year novice impact (S5)	0.041		(0.05)	0.41	**0.04**	687	1,096
(5) Two-year impact (S6)	0.086	***	(0.03)	< 0.01	**0.09**	1,113	2,651
(6) Two-year schoolwide impact (S7)	0.011		(0.02)	0.49	**0.01**	3,472	3,576
Follow-up impacts in schools that did NOT participate in the two-year study							
(7) One-year follow-up novice impact (S8)	0.063		(0.12)	0.61	**0.06**	1,209	589
(8) One-year follow-up impact (S9)	0.056	*	(0.03)	0.07	**0.06**	4,003	2,746

Notes: Seventeen districts contributed data to the student achievement impact analysis. The one-year impact analyses exclude one district, while the two-year analyses exclude two districts. Effect size is the adjusted difference between treatment and control group average standardized test scores divided by the student level unadjusted pooled standard deviation of post-test scores.

*Statistically significant at the 10 percent level, two-tailed test.

**Statistically significant at the 5 percent level, two-tailed test.

***Statistically significant at the 1 percent level, two-tailed test.

Results of this analysis show how the implementation of TPP affected the average performance level of all students in the schools at the time of testing. Findings from the analysis including student joiners showed the same pattern of results as the benchmark analysis. The impact estimates were similar in magnitude and statistical significance. The overall two-year impact (addressing research question S4) remained significant and the effect size slightly increased from 0.08 to 0.9 standard deviations, while the two-year impact effect size decreased by the same margin from 0.10 to 0.09 standard deviations. The one-year follow-up impact (addressing research question S9) effect size remained essentially the same (0.06) and retained the same level of statistical significance ($p = 0.07$). In contrast, one-year follow-up novice impact (addressing research question S8) effect size increased slightly (from 0.02 to 0.06 standard deviations. In no instance, however, did substantive findings change (compared to the benchmark model).

c. Other analyses of sensitivity of estimated impacts to alternative specifications

Additional sensitivity analyses for the following selected impacts are reported in Appendix Table G.7: the one-year impact (addressing research question S2), overall two-year impact (addressing research question S4), two-year impact (addressing research question S6), and one-year follow-up impact (addressing research question S9). The one-year impact analysis is based on the randomized controlled trial and has the largest sample size, while the two-year impact analyses

and the follow-up impact analysis all generated statistically significant findings which are important to verify against alternative model specifications.

The benchmark impact estimation model and two primary sensitivity analyses (complete case and student joiners) are presented in the first three rows of Appendix Table G.7. The alternative models displayed in other rows of the table differ from the benchmark model in the following ways:

- We used the math pre-test as an instrument for the reading pre-test to help address the possibility of measurement error in the reading measure.
- We dropped all ELA post-test scores that were more than 5 standard deviations above or below the statewide grade-specific mean to address the potential influence of outliers.
- We standardized test scores using means and standard deviations of scores within our sample of control students rather than state grade-specific means and standard deviations for specific test administered.
- We replaced district-grade interaction variables with school pair dummies.
- We replaced the control variables in the impact regression with separate district and grade indicators.
- We examined an alternative approach to addressing missing data by estimating impacts using a multiple imputation model, which involves generating multiple data sets that contain imputed values for missing data through the repeated application of an imputation algorithm.
- We examined an alternative definition of the second year of TPP using information collected from all schools in Cohort 3. In addition to the schools who implemented a second year of TPP, some schools that received only one year of TPP reported that they implemented the TPP curriculum (without the PD supports) the following year. Controlling for treatment status, we estimated impacts among Cohort 3 schools that allowed for an additional effect of a second year of the full TPP program or just the curriculum.

Overall, sensitivity tests confirm that varying assumptions do not substantially alter the magnitude and statistical significance of the main findings. In the case of the one-year impacts (addressing research question S2), each estimate is less than 0.016 and not significant. In case of the overall two-year impacts (addressing research question S4), findings are statistically significant at the 1 percent level and range from 0.048 to 0.087. Similarly, for the two-year impacts (addressing research question S6), estimated treatment coefficients vary from 0.055 to 0.101, with the lowest estimate still being significant at the 5 percent level. For the one-year follow-up impacts (addressing research question S9), findings are statistically significant at the 10 percent level for about a half of alternative model specifications. Therefore, we conclude that the main results are robust to various alternative model specifications.

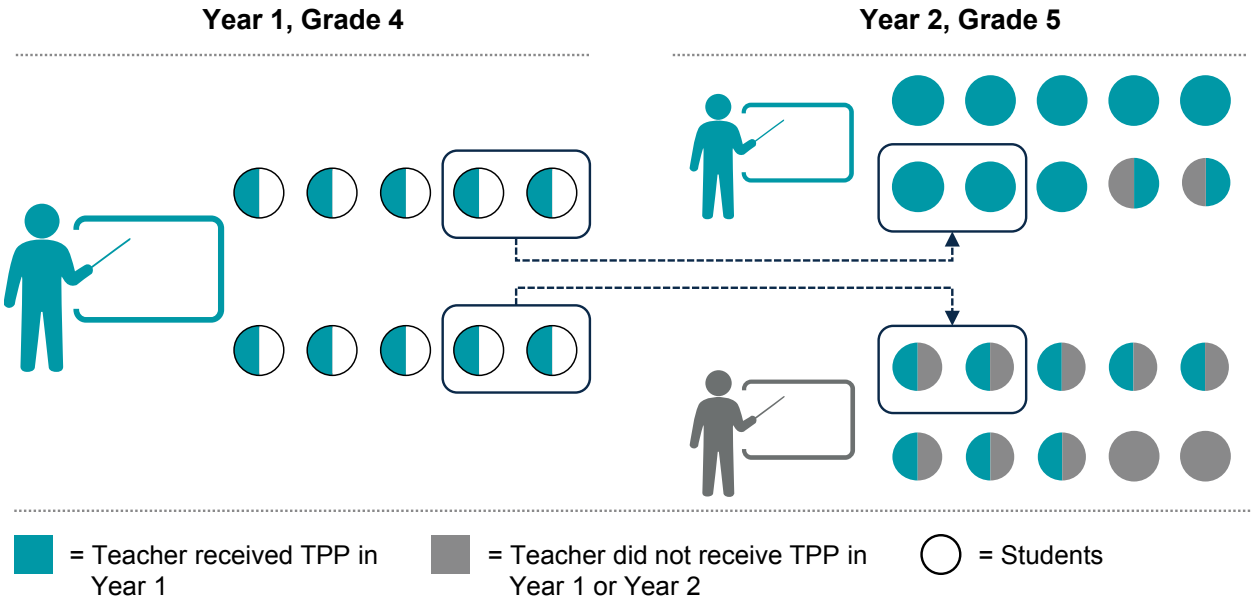
d. Considerations when interpreting the findings

There are several considerations to keep in mind when interpreting these ELA achievement analyses.

First, while the randomization of schools should lead treatment and control groups to have equivalent baseline characteristics on average, chance differences between groups are possible. For analyses involving novice teachers, we found baseline differences between treatment and comparison groups in the proportions of students at different grade levels. In particular, we found a lack of baseline equivalence between the treatment and control groups in the proportion of students in grade 8 for the one-year novice impact analysis (addressing research question S1), in grades 6 and 7 for the two-year novice impact analysis (addressing research question S5), and in grade 6 for one-year follow-up novice impact analysis (addressing research question S8) (see Appendix Table G.6). Therefore, readers should interpret student findings for novice teachers with caution.

Second, in the analyses of impacts beyond the first year, the students included in the analyses may differ in their overall exposure to TPP. Due to student and teacher mobility, students may have been taught by a TPP teacher for either one or two years. For example, consider a cohort of students who are in fourth grade in their school's first year in the study and in fifth grade during their school's second year in the study (Figure VI.1). If their school received TPP for just one year, the students would have been in the classroom of a fourth grade teacher engaging in TPP during that first year. In the second year, however, some students may be in the classroom of a fifth grade teacher who engaged in TPP the previous year, while other students may be in the classroom of a newly-hired fifth grade teacher who did not engage in TPP the previous year. Thus, during the second year, some students had been in the classroom of a TPP teacher for two years while others had been in the classroom of a TPP teacher for just one year. The same may be true for students in schools that participated in TPP for two years. The follow-up and two-year impacts presented in this chapter thus reflect an average effect during schools' second years across students who were exposed to TPP for one or two years.

Figure VI.1. Student exposure to TPP



B. Impacts on student literacy task

The CCSS includes a focus on students' use of informational texts and extended argument writing, outcomes that may not be fully captured by state standardized ELA assessments. To determine the impacts of one year of TPP implementation on student literacy (addressing research question S3), the study team engaged in an exploratory analysis of students' opinion or argument writing from spring 2017 from a subset of Cohort 3 novice ELA teachers. In this section, we first describe the literacy measures. We then describe the sample and our analytic approach to estimating impacts. Finally, we discuss the findings, including considerations when interpreting them.

1. Literacy measures

We developed a student literacy task, which consisted of an opinion writing activity (for grades 4 and 5) or an argument writing activity (for grades 6 to 8). The topic was to convince their school principal of their stance on an issue to either get a new school lunch menu (grades 4 and 5) or use video games in their class (grades 6 to 8). The literacy task was developed so the activities and outcomes would be aligned to CCSS. It was piloted in middle grades ELA classrooms in spring 2016, and it underwent two rounds of expert review to ensure the activities were clear to students and easy for teachers to implement, to establish validity, and to ensure that the activities were aligned to CCSS for the grade levels of interest. The study team asked all novice teachers in Cohort 3 to use the literacy task with all students in the first ELA class in the morning and afternoon and to collect parent consent and student assent. Two trained researchers scored the student writing on three outcomes: facets of writing conventions (such as vocabulary and grammar), facets of writing quality (such as evidence from texts and counterclaims), and general literacy, which includes all facets of writing conventions and writing quality (see Chapter 3,

Table III.1 for the description of each literacy outcome measure). The researchers analyzed the scores in aggregate by teacher. The researchers demonstrated high interrater reliability (see Appendix G).

2. Description of sample

Although the study team administered the literacy task to all novice teachers in Cohort 3 schools, not all the teachers were able to provide consent and assent and student writing samples. To preserve the matched random assignment design within each district to the extent possible in the analysis, we only included the subset of districts in which the majority of teachers in the study schools provided student writing. The analyses included all novice study teachers in all seven schools across five districts (six treatment and six control teachers). In the remaining districts, the levels of nonresponse from teachers at each school were too high to provide confidence in results drawn from comparisons between the treatment and control groups, so these schools and teachers were excluded from the analysis.

The novice teachers included in these analyses taught students in grades 5 through 8 and the majority of teachers taught grade 7. There were no grade 4 teachers in the schools and districts that were included in the analyses. Teachers were asked to use the literacy task with all their ELA classes but to provide student writing for the students in only two predetermined classes: the first ELA class in the morning and the afternoon. Approximately half the students in the teachers' two predetermined classes provided student writing samples.

The study team was unable to match student respondents to the students in the administrative data provided by the district or the students in the two predetermined classes from which the study team collected data because the districts prohibited us from collecting students' uniquely identifying information and the district data did not include information on students' ELA class period. However, the team was able to match teacher respondents to the administrative data for all their students to examine characteristics of students taught by treatment and control teachers and assess whether there were any differences between the two groups. Below, we present information on all the students taught by the respondent teachers, including the students who contributed writing samples to this study and their classmates who did not contribute writing samples.

Overall, approximately 50 percent of the students in the novice teachers' classes were white non-Hispanic, 12 percent were black non-Hispanic, and 8 percent of the students were another race (see Table VI.4). Approximately 30 percent of the students were of Hispanic ethnicity. A little more than 50 percent of the students were male. More than 60 percent qualified for free or reduced-price lunch. Approximately 15 percent of students were classified as English-language learners or as having special education needs.

Table VI.4. Comparison of characteristics of teachers' classes at baseline

	Treatment			Control					
	Mean	SD	N	Mean	SD	N	t-statistic	p-value	Effect Size
Grade	7.00	1.19	6	6.83	0.75	6	0.00	0.77	0.19
Average student standardized ELA pre-test score	–0.54	0.23	6	–0.50	0.40	6	–0.23	0.82	0.14
Percent students male	0.51	0.03	6	0.55	0.07	6	–1.22	0.25	0.69
Student race and ethnicity									
Percent black non-Hispanic	0.09	0.14	6	0.10	0.08	6	–0.24	0.82	0.14
Percent white non-Hispanic	0.53	0.19	6	0.50	0.15	6	0.31	0.76	0.19
Percent Hispanic	0.29	0.20	6	0.32	0.18	6	–0.30	0.77	0.18
Percent other race	0.09	0.08	6	0.08	0.06	6	0.43	0.68	0.26
Percent students eligible for free or reduced-price lunch	0.71	0.19	5	0.57	0.29	4	0.93	0.38	0.63
Percent students English language learner	0.09	0.05	6	0.16	0.17	6	–0.87	0.40	0.51
Percent students special education	0.12	0.10	6	0.16	0.07	6	–0.77	0.46	0.45

Source: Data are from administrative educator records for 2015–2016.

Note: Sample sizes are at the teacher level. Students' characteristics were aggregated for each teacher and then averaged within the treatment and comparison groups. Effect size is the adjusted difference between treatment and control group average student standardized test scores divided by the student level unadjusted pooled standard deviation of post-test scores.

ELA = English language arts; SD = standard deviation.

3. Attrition and baseline equivalence of treatment and control groups

As discussed in previous sections, high levels of attrition and uneven rates of attrition between the control and treatment groups can lead to the inability to compare control and treatment teachers and their students. Among teachers, there was 86 percent overall attrition and 6 percent differential attrition. Because of the high level of overall attrition, the RCT design is compromised (What Works Clearinghouse 2017). Thus, to understand the effects of TPP, we must assess whether there were baseline differences in other factors, such as students' ELA achievement, which could potentially account for differences in outcomes and lead to biased estimates of impacts. We examined whether there was baseline equivalence between treatment and control groups of teachers included in this analysis by comparing students of teachers who engaged with TPP from those who did not to test whether the groups were similar prior to TPP. We assessed baseline equivalence of teachers based on their students' average student standardized reading pre-test scores, standardized math pre-test scores, proportion male, race (white non-Hispanic, black non-Hispanic, and other), Hispanic ethnicity, free and reduced-price lunch eligibility, English language-learner status, and special education status. Across all characteristics, effect sizes were greater than 0.05, indicating that the treatment and control groups were not equivalent on all these characteristics and that any findings may be attributable to pre-existing differences between the treatment and control groups and not the TPP (What Works Clearinghouse 2017). Students of treatment teachers qualified for free or reduced-price lunches in higher proportions, and students of control teachers had lower ELA test scores and were more likely to be classified as English-language learners and special education students.

4. Analytic approach

The effects of TPP on student literacy outcomes were analyzed using a regression model that accounts for the clustering of teachers within schools. Because of the presence of group differences on baseline student characteristics, these characteristics were accounted for in the regression analyses as covariates to support the precision of the analyses and limit the bias that group differences may introduce. To do so, we conducted the analysis at the teacher level rather than the student level because we could control for average characteristics of the students in each teacher's classes but not for the characteristics of each individual student. Given the small sample size of 12 teachers, all the characteristics could not be accounted for in a single analysis. For this reason, analyses were run for three models and each included a subset of average student characteristics as specified below:

- Model 1: The first model was a regression analysis that accounted for the clustering of teachers within schools and adjusted for two characteristics: grade and the standardized ELA test scores from the prior year.

- Model 2: The second model was a regression analysis that accounted for the clustering of teachers within schools and adjusted for race (white non-Hispanic, black non-Hispanic, and other race) and ethnicity (Hispanic).

- Model 3: The third model was a regression analysis that accounted for the clustering of teachers within schools and adjusted for free or reduced-price lunch eligibility, the proportion of male students, English-language learner status, and special education status.

Due to the limited statistical power of the analyses because of the small sample sizes, the effects of TPP would need to be very large for the analyses to be able to detect any significant impacts. Findings should be interpreted with caution because the lack of statistically significant results may be driven by low power and not due to the effects of TPP.

Key finding: TPP did not have a statistically significant effect on students' writing conventions, writing quality, or overall writing scores among this subset of teachers.

5. Findings

There were no statistically significant impacts of TPP on any of the three student literacy measures (see Table VI.5).

Table VI.5. Student literacy task impacts of TPP

Literacy task outcome	Estimated treatment coefficient	Standard error	Effect size	Analytic sample size	
				Treatment	Control
Writing conventions domain					
Vocabulary	–0.07	0.31	–0.76	6	6
Mechanics	0.05	0.26	0.16	6	6
Spelling	0.01	0.10	–0.01	6	6
Grammar	–0.02	0.24	0.02	6	6
Writing conventions total	–0.02	0.74	–0.01	6	6
Writing quality domain					
Voice	–0.30	0.10	–2.01	6	6
Sentence variety	–0.34	0.13	–3.87	6	6
Organization	–0.33	0.35	–1.28	6	6
Statement of position	–0.03	0.09	–0.27	6	6
Evidence	0.18	0.22	0.82	6	6
Reasons	–0.11	0.31	–0.11	6	6
Transition words and links	0.10	0.28	0.59	6	6
Counterclaims (grades 7–8)	–0.14	0.32	–1.41	5	4
Writing quality total	–0.19	1.00	–0.09	6	6
Writing quality (grades 7–8)	–0.47	1.13	–0.50	6	6
General literacy domain					
Total score	–0.91	1.73	–0.36	6	6
Total score (grades 7–8)	–1.00	3.61	–0.87	5	4

Source: Impacts were estimated on an average of students' writing scores for novice teachers in a subsample of schools, accounting for the prior year's standardized ELA test as the baseline measure.

Note: Effect size is the adjusted difference between treatment and control group average student standardized test scores divided by the unadjusted pooled standard deviation with an adjustment for small sample sizes per What Works Clearinghouse guidance.

Results are presented for Model 1. There were no significant results for Models 2 and 3. No findings were statistically significant with multiple comparison corrections.

Writing conventions. When controlling for the effect of student grade and ELA pre-test scores (Model 1), there were no effects of TPP on student writing conventions overall or for vocabulary, mechanics, spelling, and grammar subscores.

Similarly, there were no effects of TPP on student writing conventions when controlling for race and ethnicity (Model 2) and no effects of TPP on student writing conventions when controlling for free or reduced-price lunch eligibility, proportion of male students, English-language learner status, and special education status (Model 3).

Writing quality. There were no significant differences between treatment and control groups after adjusting for the effect of student grade and ELA pre-test scores on overall writing quality or on any of the facets of writing quality, such as organization and statement of position (Model 1). Similar results held for Models 2 and 3.

General literacy. After controlling for student grade and ELA pre-test score, there were no significant differences between treatment and control groups on the total writing score for all grades and for grades 7 and 8, which included a score for counterclaims (Model 1). Similar results held for Models 2 and 3.

6. Considerations when interpreting the findings

Through this exploratory analysis, we examined differences between a sampling of treatment and control teachers' students on literacy outcomes. Although these analyses provided some insight into the effect of TPP on student writing, there are some key considerations for readers when interpreting the findings.

We were limited to including only a small subset of the districts, schools, and teachers in the broader study because of high levels of nonresponse at the teacher level. This resulted in underpowered analyses of 12 teachers and a disruption to our RCT design. With such a small sample size, the study was underpowered to detect differences among groups and further research is needed to understand the role of TPP on students' writing conventions and opinion or argument writing quality.

VII. DISCUSSION

Integrating teacher professional learning with educative curriculum has emerged as a promising way to both build teacher capacity and support student achievement. Studies have found that much current professional development does not meet teachers' needs (Bill & Melinda Gates Foundation 2014; Calvert 2016) and may not be associated with improvement in teacher performance (TNTP 2015). One key reason may be that some supports for instructional practice are not based in teachers' specific context (Taylor et al. 2015, Wiener and Pimentel 2017). This puts more demand on teachers to adapt guidance on practices to their own classroom and curriculum (Stein et al. 2007). In contrast, providing teachers with specific strategies and support for the material they are delivering can make professional development more relevant and easy to use at once. An analysis of 35 studies of PD concluded that it is most effective when it is focused on "specific curriculum content" (Darling-Hammond et al. 2017).

This type of support embedded in curriculum can be especially important for novice teachers who may be gaining familiarity with the curriculum and school environment at the same time as they are gaining experience with instructional practices. Most districts offer novices some form of mentoring or induction services (defined as support, guidance, and orientation programs; see Berry et al. 2002 and Smith and Ingersoll 2004), but these are not often comprehensive or systematic (Kauffman et al. 2002) and are often designed to address teacher retention rather than teacher practice (Glazerman et al. 2010). Moreover, there is little evidence that induction affects novice teacher practices (Lopez et al. 2004; Glazerman et al. 2010).

TPP, and the innovative approach developed by EL Education on which TPP is based, is one of the few programs providing curriculum with embedded PD. TPP focuses on ELA curriculum and professional learning supports for teachers in upper elementary and middle school grades. Mathematica designed an RCT and a two-year QED to evaluate the impacts of TPP on students and teachers across 18 school districts nationwide. Results presented in this report show that TPP is effective in improving student achievement and teacher instructional practice outcomes.

In this chapter, we synthesize the findings of this evaluation to better understand the changes brought about by TPP. We begin with a summary of the findings on implementation, teacher impacts, and student impacts. Next, we discuss the patterns and size of the findings in the context of recent literature. Finally, we discuss potential explanations for our findings.

A. Synthesis and interpretation of findings

This evaluation found the following impacts of TPP after one and two years, on novice and all teachers and on students.

Findings after one year. Teachers who engaged in the first year of TPP generally participated in program services as intended (except in their use of online supports). TPP led to improvements in teachers' overall ELA instructional practices, as well as specific practice outcomes, during that year. The specific practices that were impacted were in the topics of having connected

lessons; supporting students' higher-order thinking; engaging students in reading, writing, and/or speaking about texts; supporting students' use of text evidence; and supporting student participation. There were no effects on their students' achievement, either on standardized ELA assessments or on a literacy task for students of novice teachers, at the end of the first year. However, for teachers who only engaged in a single year of TPP, there were positive effects approaching the 5 percent level of significance on the achievement of their students at the end of the following year.

Findings after two years. Teachers who engaged in two years of TPP participated in program services during the second year as intended (except in their use of online supports). There were no impacts on the instructional practice constructs during these teachers' second year of engagement, likely due to a smaller sample size for these analyses, which would inhibit the analyses' ability to fully detect impacts of TPP. There was a positive and significant impact on the achievement of students in these teachers' classrooms in the second year, with an effect size of 0.10 standard deviations. This effect size of TPP can be interpreted as moving an average student scoring at the 50th percentile to the 54th percentile. It is equivalent to roughly 1.4 months of typical student improvement (assuming 0.7 standard deviations of improvement over a 10-month school year). These impacts were only among the students of teachers who engaged with TPP directly and did not appear to affect other nonstudy ELA teachers. They do not appear to be due to a more proficient sample of teachers in these selected schools, but rather can be attributed to TPP.

However, student impacts in schools' second study year were observed more broadly than just among schools participating in two years of TPP implementation. Among all schools that engaged with TPP—for a single year or for two years of implementation—there was a positive and significant impact on student achievement in the second year with an effect size of 0.08 standard deviations. This effect size of TPP can be interpreted as moving an average student scoring at the 50th percentile to the 53th percentile, or as equivalent to roughly 1.1 months of typical student improvement.

Findings for novice teachers. Novice teachers participated in program services as intended during the first and second years of TPP implementation and experienced positive impacts in the same practices as all study teachers. These impacts on novice practice were generally larger in size than among all study teachers. There was one exception: in their first year of engagement with TPP, there was no impact on practices supporting student participation among novice teachers. However, there were no significant impacts on the achievement of students in the classrooms of novice teachers during their first or second year of engagement with TPP.

B. Discussion of findings in the context of related literature

Programs pairing professional development with educative curriculum have received increasing attention in the recent literature. This study of TPP makes several important contributions to this literature. It evaluates a comprehensive intervention in a broad curricular area—English language arts—and examines the intervention's effects on accordingly broad outcomes of ELA instruction and achievement; this is unlike a number of other studies that examine programs and outcomes

that are focused on more targeted topic areas. Additionally, this study uses rigorous group designs to identify causal impacts of the intervention and examines both one and two years of implementation. Lastly, the study features both teacher and student outcomes to examine the effects of the intervention on teachers, who engage firsthand in the intervention, and students, who may benefit from teachers' TPP participation.

To help illustrate our contribution and provide a context to better understand our findings, in this section we discuss the results of several studies from the literature evaluating paired PD and curriculum programs. These studies have examined such programs in different grades and subject areas and investigated their impacts on teachers and on students. We focus on the limited number of studies that are similar to ours in using rigorous group designs (RCT and QED) to assess programs implemented for one or two years. Studies of a few similar paired PD and curriculum programs such as Eureka Math™ and Wit & Wisdom® that examined case studies are not discussed.

Only three studies that we identified examined two-year programs of paired PD and curriculum as this report does for the two-year QED study. All found significant impacts on student achievement. Gallagher (2017) conducted a two-year RCT evaluation of the National Writing Project's College-Ready Writers Program that serves secondary school teachers and students. The program significantly improved student performance on measures of writing content, structure, stance, and conventions, with effect sizes ranging from 0.12 to 0.2. Johnson and Fargo (2010) conducted a two-year QED evaluation of a middle school science program combining PD and curriculum. Student performance on a specially designed science assessment was significantly higher among treatment students at the end of the second year (effect size not reported). While both sets of impacts are positive, they examine rather narrowly-defined outcomes. Newman et al. (2012) examines a more general measure of student achievement in science. The authors conducted a two-year RCT evaluation of the Alabama Math, Science, and Technology Initiative in grades 4 through 8. The program had positive and significant impacts on students, with an effect size of 0.10 on math test scores and 0.13 on science test scores. These impacts on writing, science, and math are similar in size to the impacts we found on ELA achievement.

All three of these two-year studies found impacts on teacher practice as well, though these impacts are on specific practices rather than more general measures of overall standards-aligned instruction. Gallagher et al. (2017) found impacts on elements of argument writing emphasized by the program, including developing a claim, evaluating evidence that could be used in support of the claim, developing an argument in support of the claim, and practicing the argument. Johnson and Fargo (2010) found impacts on school-level design and implementation of the lesson. Newman et al. (2012) found impacts on teachers spending more time using active learning instructional strategies.

A larger set of studies examined one-year programs that included both PD and curriculum components (Kemple 2008; James-Burdumy 2009; Buysse et al. 2010; Doppelt 2009; Finkelstein 2010; Kleickmann et al. 2016; Kutaka et al. 2017; Landry et al. 2009; Lara-Alecio et al. 2012; May et al. 2016; Saxe et al. 2001). These programs spanned a range of subject areas, grades, and

special student populations, such as English-language learners and struggling readers. One of these studies examined teacher practice outcomes and found impacts on language and literacy practices (Buysse et al. 2010). All examined student outcomes, and all but one found positive and significant impacts on students' achievement. The studies examined a variety of student achievement outcomes on a variety of assessment instruments. Outcomes included test scores in the areas of math, science, reading, early childhood literacy, and economic literacy. Impact estimates spanned a wide range and were often large in size, although many of the largest impacts were for programs targeting special populations and for programs assessed using very specific outcome measures that may have been closely tied to specific skills and objectives of the curriculum or instructional strategies (for example, rhyme matching and knowledge of scientific principles of floating and sinking).

C. Potential explanations for findings

This study was not designed to identify the mechanisms by which TPP achieved impacts on teachers and students. We offer, instead, some potential explanations that would be consistent with the patterns of findings but present these as no more than hypotheses. Future work to investigate these hypotheses could advance our understanding still further.

The patterns of findings over the first and second years of the study show that, while impacts on teacher practices appeared during the first year of engagement with TPP, impacts on students did not appear until after the first year. What could account for this lag in student impacts? It would be consistent with an explanation that teachers may need time to fully digest and implement the supports that they receive (Glazerman et al. 2010, Isenberg et al. 2010). Even though impacts on practices are evident during the first year, there may be changes we are unable to observe that come about in the second year once teachers have met their initial "survival" needs. Thus, this explanation requires that there are limitations to our observed measures of practices.

A second explanation is that it is possible that a longer cumulative exposure of teachers to PD is needed to affect student achievement. A review of the literature on the effects of PD on student achievement found that substantial time spent in PD was needed to lead to effects on students (Yoon et al. 2007). This explanation is also consistent with our finding that student impacts are larger in the second year than in the first year, as teachers build up more PD exposure.

Third, it is possible that sustained PD is also important. A review of 35 rigorous studies of professional development programs found that PD was most effective when it involved sustained, ongoing engagement, with substantial time spent and multiple learning opportunities offered (Darling-Hammond et al. 2017). This explanation is also consistent with our finding that student impacts are larger in the second year than in the first year, but is not necessarily consistent with our finding of positive follow-up impacts that approach statistical significance.

There are other possible explanations for the lag in student impacts. A fourth possible explanation is that the changes in teachers' instructional practice during the first year were not large enough to affect student achievement that year; studies have found that impacts on teachers' practice must be large to impact students (Quint 2011). A fifth explanation could relate

to the timing of changes in teachers' instructional practice; it may be that impacts on teachers' practice were not achieved early enough in the school year for students to see the benefits of the changes in teachers' instructional practices.

It is also notable that the study did not find impacts on teacher practices during their second year of engagement with TPP. While it is possible that TPP did not affect practices during the second year, another potential explanation for this result is that this study was limited in its ability to detect impacts in teachers' instructional practices after teachers engaged with TPP in two years. Prior study findings on TPP's impacts on teachers' instructional practices suggest that there are differences between treatment and control groups on specific instructional practices after two years of teacher participation in TPP; these differences included treatment teachers having students explain or support their understanding of what they had read or cite evidence from the text to support their responses verbally or in writing more than control teachers (Choi et al. 2018). When examining this earlier finding alongside the finding in the current report of the impact of two years of TPP on students' ELA achievement, they could mean that TPP did have an impact on teachers' instructional practices in their second year that this study could not capture for several possible reasons. One reason is related to the small numbers of teachers included in the two year sample in this study's analysis approach. Another reason is that the broader constructs of teachers' instructional practices used in this report did not adequately reflect some of the specific, nuanced differences between treatment and control teachers' practices covered in the earlier study findings. Still another reason is that the study team developed the instructional practice constructs using data from teachers' first year with TPP and applied these same constructs to measure teacher's instructional practices in their second year (after checking model fit to ensure that the constructs adequately captured both years of data). The same constructs were used for both years of data in order to have consistently defined measures across the study's different samples of teachers over the two years. However, this may mean that the constructs were more reflective of teachers' instructional practices in their first year than their second year, and did not capture some of the differences between treatment and control teachers' practices that could have impacted student achievement in the second year.

Yet another notable pattern of findings was the lack of impacts on novice teachers' students, despite impacts on these teachers' instructional practices. This pattern is consistent with an explanation that it takes time for professional learning support to lead to impacts on students. It is possible that it takes several years for support to impact novices' classrooms (Glazerman et al. 2010)—longer than for teachers on average since novices are starting at a baseline of less professional experience. Some evidence from this evaluation supports this idea. Coaches stated in their surveys that some novice teachers struggled with classroom management. It may have been challenging for novice teachers to juggle learning and integrating classroom management strategies in addition to other instructional strategies and practices.

This page has been left blank for double-sided copying.

REFERENCES

Allen, Joseph P., Robert C. Pianta, Anne Gregory, Amori. Y. 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. doi:10.1126/science.1207998.

Amoruso, M., B. D. Bontempo, and D. J. Wilson. "The Relationship Between ELS Participation and Academic Growth." Portland, OR: Mountain Measurement, Inc., 2010.

Atkinson, Thomas M., Barry D. Rosenfeld, Laura Sit, Tito R. Mendoza, Mike Fruscione, Dawn Lavene, Mary Shaw, Yuelin Li, Jennifer Hay, Charles S. Cleeland, Howard I. Scher, William S. Breitbart, and Ethan Basch. "Using Confirmatory Factor Analysis to Evaluate Construct Validity of the Brief Pain Inventory (BPI)." *Journal of Pain and Symptom Management*, vol. 41, no. 3, 2011, pp. 558–565. doi:10.1016/j.jpainsymman.2010.05.008.

Baker, Scott, Nonie Lesaux, Madhavi Jayanthi, Joseph Dimino, C. P. Proctor, Joan Morris, Russel Gersten, Kelly Haymond, Michael J. Kieffer, Sylvia Linan-Thompson, and Rebecca Newman-Gonchar. "Teaching Academic Content and Literacy to English Learners in Elementary and Middle School." NCEE 2014-4012. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, April 2014.

Benjamini, Yoav and Yosef Hochberg. (1995). "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society*. Series B (Methodological), vol. 57, no. 1, 1995, pp. 289–300. doi:10.2307/2346101.

Berry, Barnett, Peggy Hopkins-Thompson, and Mandy Hoke. "Assessing and Supporting New Teachers: Lessons from the Southeast." Chapel Hill, NC: Southeast Center for Teaching Quality, December 2002.

Bill & Melinda Gates Foundation. "Teachers Know Best: Teachers' Views on Professional Development." Seattle, WA: The Bill & Melinda Gates Foundation, December 2014.

Borman, Geoffrey D., Samuel C. Stringfield, and Robert E. Slavin (eds.). *Title I: Compensatory Education at the Crossroads.* Mahwah, NJ: Erlbaum, 2001.

Buysse, Virginia, Dina C. Castro, and Ellen Peisner-Feinberg. "Effects of a Professional Development Program on Classroom Practices and Outcomes for Latino Dual Language Learners." *Early Childhood Research Quarterly*, vol. 25, 2010, pp. 194–206. doi:10.1016/j.ecresq.2009.10.001.

Calvert, Laurie, "Moving from Compliance to Agency: What Teachers Need to Make Professional Learning Work." Oxford, OH: Learning Forward and NCTAF, March 2016.

Choi, Jane, Scott Richman, and Sarah Dolfin. "Transforming Teachers' Practice: The Impact of EL Education's English Language Arts Curriculum and Professional Learning on Teacher Practices." Princeton, NJ: Mathematica Policy Research, October 2017.

Choi, Jane, Scott Richman, and Sarah Dolfin. "Sustaining Effective Teacher Practice: The Impact of the EL Education Language Arts Curriculum and Professional Development on Teachers' Instruction." Princeton, NJ: Mathematica Policy Research, November 2018.

Cohen, Jacob. *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Mahwah, NJ: Erlbaum, 1988.

Darling-Hammond, Linda, Maria E. Hyler, and Madelyn Gardner. "Effective Teacher Professional Development." Palo Alto, CA: Learning Policy Institute, 2017.

Doppelt, Yaron, Christian D. Schunn, Eli M. Silk, Matthew M. Mehalik, Birdy Reynolds, and Erin Ward. "Evaluating the Impact of a Facilitated Learning Community Approach to Professional Development on Teacher Practice and Student Achievement." *Research in Science & Technological Education*, vol. 27, no. 3, November 2009, pp. 339–354. doi:10.1080/02635140903166026.

Isenberg, Eric, Steven Glazerman, Amy Johnson, Sarah Dolfin, and Martha Bleeker. "Linking Induction to Student Achievement." In *Past, Present, and Future Research on Teacher Induction: An Anthology for Researchers, Policy Makers, and Practitioners*, edited by J. Wang, S.J. Odell, and R.T. Clift. Lanham, MD: Rowman and Littlefield Education, June 2010.

Finkelstein, Neal, Thomas Hanson, Chun-Wei Huang, Becca Hirschman, and Min Huang. "Effects of Problem Based Economics on High School Economics Instruction." NCEE 2010-4002. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, July 2010.

Gadermann, Anne M., Martin Guhn, and Bruno D. Zumbo. "Estimating Ordinal Reliability for Likert-type and Ordinal Item Response Data: A Conceptual, Empirical, and Practical Guide." *Practical Assessment, Research & Evaluation*, vol. 17, no. 3, 2012, pp. 1–13.

Gallagher, Alix H., Nicole Arshan, and Katrina Woodworth. "Impact of the National Writing Project's College-Ready Writers Program in High-Need Rural Districts." *Journal of Research on Educational Effectiveness*, vol. 10, no. 3, 2017, pp. 570–595. doi:10.1080/19345747.2017.1300361.

Glazerman, Steven, Eric Isenberg, Sarah Dolfin, Martha Bleeker, Amy Johnson, Mary Grider, and Matthew Jacobus. "Impacts of Comprehensive Teacher Induction: Final Results from a Randomized Controlled Study." NCEE 2010-4027. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, June 2010. https://www.mathematica-mpr.com/our-publications-and-findings/projects/supporting-new-teachers-evaluating-teacher-induction-models.

Hox, Joop. *Multilevel Analysis: Techniques and Applications* (2nd ed.). New York, NY: Routledge Academics, 2010.

James-Burdumy, Susanne, Wendy Mansfield, John Deke, Nancy Carey, Julieta Lugo-Gil, Alan Hershey, Aaron Douglas, Russell Gersten, Rebecca Newman-Gonchar, Joseph Dimino, and Bonnie Faddis. "Effectiveness of Selected Supplemental Reading Comprehension Interventions: Impacts on a First Cohort of Fifth-Grade Students." NCEE 2009-4032. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, May 2009.

Johnson, Carla C., and Jamison D. Fargo. "Urban School Reform Enabled by Transformative Professional Development: Impact on Teacher Change and Student Learning of Science." *Urban Education*, vol. 45, no. 1, 2010, pp. 4–29. doi: 10.1177/0042085909352073.

Kauffman, David, Susan Moore Johnson, Susan M. Kardos, Edward Liu, and Heather G. Peske. "Lost at Sea: New Teachers' Experiences with Curriculum and Assessment." *Teachers College Record*, vol. 104, no. 2, 2002, pp. 273–300.

Kemple, James J., William Corrin, Elizabeth Nelson, Terry Salinger, Suzannah Herrmann, and Kathryn Drummond. "The Enhanced Reading Opportunities Study: Early Impact and Implementation Findings." NCEE 2008-4015. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, January 2008.

Kleickmann, Thilo, Steffen Tröbst, Angela Jonen, Julia Vehmeyer, and Kornelia Möller. "The Effects of Expert Scaffolding in Elementary Science Professional Development on Teachers' Beliefs and Motivations, Instructional Practices, and Student Achievement." *Journal of Educational Psychology*, vol. 108, no. 1, 2016, pp. 21–42. doi:10.1037/edu0000041.

Kutaka, Traci Shizu, Wendy M. Smith, Anthony D. Albano, Carolyn Pope Edwards, Lixin Ren, Heidi Lynn Beattie, W. James Lewis, Ruth M. Heaton, and Walter W. Stroup. "Connecting Teacher Professional Development and Student Mathematics Achievement: A 4-Year Study of an Elementary Mathematics Specialist Program." *Journal of Teacher Education*, vol. 68, no. 2, 2017, pp. 140–154. doi:10.1177/0022487116687551.

Landry, Susan H., Jason L. Anthony, Paul R. Swank, and Pauline Monseque-Bailey. "Effectiveness of Comprehensive Professional Development for Teachers of At-Risk Preschoolers." *Journal of Educational Psychology*, vol. 101, no. 2, 2009, pp. 448–465. doi:10.1037/a0013842.

Lara-Alecio, Rafael, Fuhui Tong, Beverly J. Irby, Cindy Guerrero, Maggie Huerta, and Yinan Fan. "The Effect of an Instructional Intervention on Middle School English Learners' Science and English Reading Achievement." *Journal of Research in Science Teaching*, vol. 49, no. 8, 2012, pp. 987–1011. doi:10.1002/tea.21031.

Lopez, Alejandra, Andrea Lash, Monika Schaffner, Patrick Shields, and Mary Wagner. "Review of Research on the Impact of Beginning Teacher Induction on Teacher Quality and Retention." Menlo Park, CA: SRI International, February 2004.

May, Henry, Philip M. Sirinides, Abigail Gray, and Heather Goldsworthy. "Reading Recovery: An Evaluation of the Four-Year i3 Scale-Up." Philadelphia, PA: Consortium for Policy Research in Education, 2016.

Muthén, Bengt, and Tihomir Asparouhov. "Bayesian SEM : A More Flexible Representation of Substantive Theory." *Psychological Methods*, vol. 17, no. 3, 2012, pp. 313–335. doi:10.1037/a0026802.

Newman, Denis, Pamela B. Finney, Steve Bel, Herb Turner, Andrew P. Jaciw, Jenna L. Zacamy, and Laura F. Gould. "Evaluation of the Effectiveness of the Alabama Math, Science, and Technology Initiative (AMSTI)." Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education, February 2012.

Nichols-Barrer, Ira, and Joshua Haimson. "Impacts of Five Expeditionary Learning Middle Schools on Academic Achievement." Princeton, NJ: Mathematica Policy Research, July 2013.

Park, Jungkyu, and Hsiu-Ting Yu. "Recommendations on the Sample Sizes for Multilevel Latent Class Models." *Educational and Psychological Measurement*, vol. 78, no. 5, 2018, pp. 737–761. doi:10.1177/0013164417719111.

Quint, Janet. "Professional Development for Teachers: What Two Rigorous Studies Tell Us." New York, NY: MDRC, July 2011.

Saxe, Geoffrey, Maryl Gearhart, and. Na'ilah Suad Nasir. "Enhancing Students' Understanding of Mathematics: A Study of Three Contrasting Approaches to Professional Support." *Journal of Mathematics Teacher Education*, vol. 4, issue 1, January 2001, pp. 55–79.

Schmidt, Rebecca, Viki Young, Lauren Cassidy, Haiwen Wang, and Katrina Laguarda. "Impact of the New Teacher Center's New Teacher Induction Model on Teachers and Students." Menlo Park, CA: SRI International, June 2017.

Smith, Thomas M., and Richard M. Ingersoll. "What Are the Effects of Induction and Mentoring on Beginning Teacher Turnover?" American Educational Research Journal, vol. 41, no. 3, 2004, pp. 681-714.

Stein, Mary K., Janine Remillard, and Margaret Smith. "How Curriculum Influences Student Learning." In *Second Handbook of Research on Mathematics Teaching and Learning,* edited by Frank K. Lester, Jr. (pp. 319–369). Charlotte, NC: Information Age, 2007.

Tabachnik, Barbara G. and Linda S. Fidell. *Using Multivariate Statistics* (4th ed.). Boston, MA: Allyn & Bacon, 2001.

Taylor, Joseph. A., Stephen. R. Getty, Susan M. Kowalski, Christopher D. Wilson, Janet Carlson, and Pamela Van Scotter. "An Efficacy Trial of Research-Based Curriculum Materials with Curriculum-Based Professional Development." *American Educational Research Journal*, vol. 52, no. 5, pp. 984–1017, 2015.

TNTP. "The Mirage: Confronting the Hard Truth about Our Quest for Teacher Development." New York, NY: TNTP, 2015.

UMass Donahue Institute, Research and Evaluation Group. "Impact of the Expeditionary Learning Model on Student Academic Performance in Rochester, New York." Hadley, MA: UMass Donahue Institute, Research and Evaluation Group, July 2011.

van Erp, Sara, Joris Mulder, and Daniel L. Oberski. "Prior Sensitivity Analysis in Default Bayesian Structural Equation Modeling." *Psychological Methods*, vol. 23, no. 2, 2018, pp. 363–388. doi:10.1037/met0000162.

What Works Clearinghouse. *Standards Handbook Version 4.0*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2017.

Wiener, Ross, and Susan Pimentel. "Practice What You Teach: Connecting Curriculum & Professional Learning in Schools." Washington, DC: The Aspen Institute, April 2017.

Yoon, Kwank Suk, Teresa Duncan, Silvia Wen-Yu Lee, Beth Scarloss, and Kathy L. Shapley. "Reviewing the Evidence on How Teacher Professional Development Affects Student Achievement." REL 2007–No. 033. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest, 2007.

Zinbarg, Richard E., William Revelle, Iftah Yovel, and Wen Li. "Cronbach's σ, Revelle's β, and McDonald's ω: Their relations with Each Other and Two Alternative Conceptualizations of Reliability." *Psychometrika*, vol. 70, no. 1, 2005, pp. 123–133. doi:10.1007/s11336-003-0974-7.

This page has been left blank for double-sided copying.

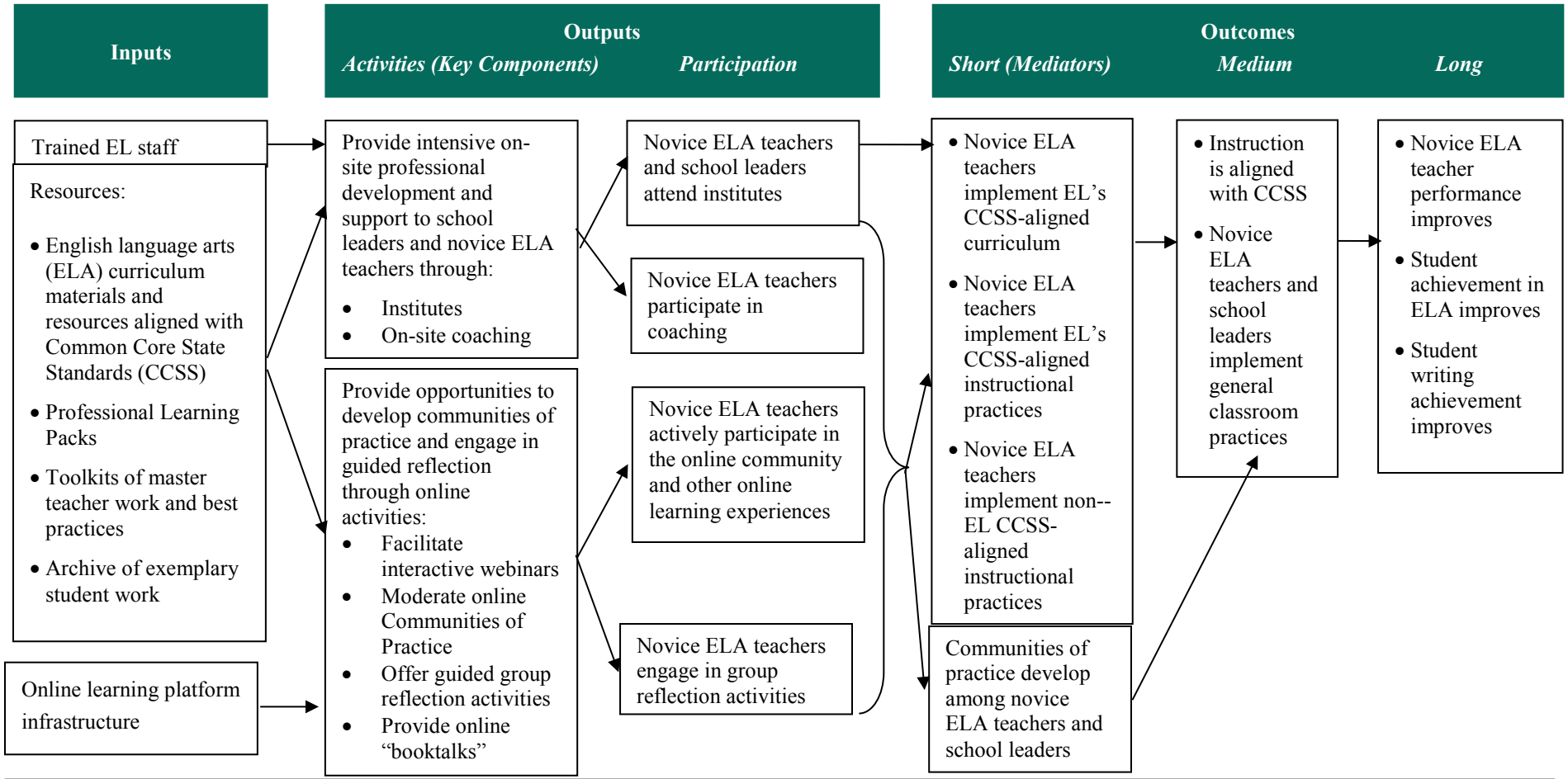
Appendix A:

Study logic model

This page has been left blank for double-sided copying.

This appendix presents a logic model showing the theory of action of TPP. The logic model forms the basis for this study. The logic model reflects the primary focus of TPP and the evaluation of novice teachers. However, it also applies to teachers in general, regardless of their level of experience.

This page has been left blank for double-sided copying.



Necessary Conditions

Schools and districts:

- Provide adequate team planning time for novice teachers to implement the curriculum.
- Release novice teachers for all necessary professional development.
- Provide necessary technology for participants to access online components of the professional development.
- Provide at least 45 minutes daily for ELA instruction at the middle level and at least 60 minutes daily for elementary ELA instruction.
- Develop and monitor a work plan in conjunction with EL Education staff

This page has been left blank for double-sided copying.

Appendix B:

Method of random assignment

This page has been left blank for double-sided copying.

Schools were randomly assigned to treatment or control conditions within each district. Within districts, random assignment was conducted among matched pairs of schools. Elementary and middle schools were paired separately. All schools had a 50 percent chance of being assigned to each condition. This appendix describes the method of matching and random assignment.

Schools were matched based on school-level characteristics available from several sources. First, we used information on schools' average English language arts and math performance from spring 2013 from GreatSchools.org to ensure that the pairs would be similar in terms of baseline student achievement. Next, we used information on school grade configuration, charter status, student enrollment and demographics, number of teachers, and student-teacher ratio from the latest year of Common Core of Data (CCD) available at the time of random assignment (from 2011–2012, 2012–2013, or 2013–2014 depending on the district), as well as the percentage of students who were English-language learners from GreatSchools.org (from 2013–2014 or 2014–2015), to ensure that pairs had a similar size and student body background. Finally, we asked principals of study schools for the number of novice ELA teachers they expected to have at the school (existing novices plus expected new hires) in the coming fall, to ensure that pairs would be similar in terms of the sample size of novice teachers in our study. Table B.1 summarizes the information available and their sources.

We developed a matching model, which refers to a set of matching variables, for each district separately. We used combinations of the following variables:

- School charter status
- Percentage of students eligible for free and reduced-price lunch
- Number of novice ELA teachers expected in the coming school year
- Percentage of students reaching proficiency in ELA and math
- An indicator related to grade configuration and/or number of students (for example, whether any grade 8 students were in the school, sometimes together with total number of students, or number of grade 4 students)
- A set of race/ethnicity indicators for groups that had more than 10 percent representation across the schools
- Sometimes pupil-teacher ratio, number of full-time equivalent teachers, and/or percent of students female

We standardized these variables using state means and standard deviations, or sample means and standard deviations if state statistics were not available, to avoid having large values swamp the estimates. We explored variations in the set of explanatory variables included with the goal of improving the match quality. We aimed to include similar numbers of variables in different domains so that one domain would not implicitly receive extra weight in matching. For example, we did not include multiple measures of the number of students in the school or each grade.

We matched schools into pairs within each district as follows. Using a set of matching variables, we computed the Mahalanobis distance—the distance between two points in multivariate space—between school pairs, for each possible grouping of schools into pairs. We then summed functions of the Mahalanobis distance between all the pairs in each possible grouping. We examined three such sums or distance metrics: the sum of squared distances, the sum of absolute distances, and the sum of variable order rankings. We examined the groupings of sets of pairs to identify the set of pairs that minimized the distance metrics. We selected the set of pairs that minimized all or most of the distance metrics.

Within the selected set of school pairs in each district, we randomly assigned one school in each pair to the treatment condition and the other to the control condition. When districts had an odd number of elementary or middle schools, there would be one "singleton" school that could not be paired. To determine an experimental assignment for these cases, we generated two lists of experimental assignments: one for elementary and one for middle schools. The lists contained treatment and control assignments randomly ordered for every two items. Whenever we obtained a singleton elementary or middle school in the random assignment process, we gave it the next assignment on the appropriate list. This procedure ensured that these schools would be randomly assigned to one of the two groups and that there would be a similar number of treatment and control schools overall (either equal or different by one).

In districts that had only two schools, we would randomly assign them to treatment and control with equal probability without matching. In two states, one in Cohort 2 and one in Cohort 3, we had two districts that each identified one school for inclusion in the study. Within each of these states, we randomly assigned one district to the treatment condition and one to the control condition with equal probability without matching.

Table B.1. School-level data items used for matching

Variable	Source
Percentage reaching language arts proficiency, spring 2013	Greatschools.org
Percentage reaching math proficiency, spring 2013	
Percentage of students who are limited English proficient	
Grade configuration (for example, K–8)	Common Core of Data
Total student enrollment	
Percentage of female students	
Percentage of students in race/ethnicity groups	
Percentage of students eligible for free/reduced-price lunch	
Student/teacher ratio	
Number of expected novice teachers in the fall	Principals of study schools

Appendix C:

Teacher survey

This page has been left blank for double-sided copying.

This appendix contains the teacher survey described in Chapter III. The survey was fielded four times: fall 2015, spring 2016, fall 2016, and spring 2017. Several changes were made each round, as summarized below. The instrument shown here is the final version used in spring 2017.

- Each survey cycle, the reference date was changed from June 20XX (in the fall) to January 20XX (in the spring).
- Before the fall 2015 and spring 2016 surveys, several formatting and minor question wording changes were made.
- Before the spring 2016 survey, several new questions were added to replace earlier "other" category or "specify" questions, or both, in the following items: A2l and A3l, B1, B3, B5, D6, E12, G1f1, G1f2, and G1f3.
- Before the spring 2016 survey, two new questions G3a and G3b were added to collect information on subjects and grade levels taught, and question G9 was redesigned to collect more accurate information on the amount of time teaching. Response categories d, e, and f were removed from G9 and G9a, and G9b was created to collect that information.
- Before the fall 2016 survey, new questions E15 to E20 were added to collect information on TPP's perceived influence on teachers' approaches to teaching and teacher understanding of state standards.

MATHEMATICA
Policy Research

EL EDUCATION
TEACHER POTENTIAL PROJECT
[REFERENCE DATE] TEACHER SURVEY

PLEASE READ THIS ENTIRE PAGE BEFORE COMPLETING THE SURVEY

Welcome to the EL Education Teacher Potential Project
[REFERENCE DATE] Teacher Survey

EL Education and Mathematica Policy Research are conducting a study to explore teachers' experiences implementing the new Common Core State Standards in English Language Arts (ELA). To learn about teachers' instructional practices, we are conducting a web-based survey of all ELA teachers and classroom observations of selected teachers in the winter and spring of the 2015-2016 school year.

Your participation is important. By completing this survey you are providing information that will help the study team: (1) identify the range of instructional practices used in ELA classrooms; (2) identify the diversity of teacher professional development activities in your school; and (3) help EL Education gauge the success of its program and improve its program (for those teachers in schools where EL Education's program is being implemented).

All of the information you provide will be treated as confidential. This survey is being conducted by Mathematica Policy Research, an organization experienced in administering surveys to teachers and in following federal guidelines to ensure confidentiality. All of the information you provide will be used only in aggregated form, so that it will not be possible to discern the identity of any survey participant in any report or presentation, or in any public use files that may be made available to researchers at the conclusion of this study. (The survey uses secure web-based technology and any information that could be used to identify or link responses to an individual respondent for any survey question will be maintained in storage that is secure.)

Your participation is voluntary. Completing the survey takes approximately 30 minutes. You may choose not to answer any question or discontinue participation at any point. There is no personal risk to you in responding to this survey. Your identity will be known only to Mathematica and EL Education staff. No information concerning respondents (including who has agreed to participate) will be given to your school. **Please take the time to complete the survey within the next two weeks.**

If you have questions about the study or the survey or if you are having technical difficulties, please email Mathematica at ELTeacherSurvey@mathematica-mpr.com or call toll free at 1-844-376-8159.

☐ **Click here to indicate that you consent to participate in this study.**

A. PROFESSIONAL DEVELOPMENT

This section includes questions about all English/Language Arts (ELA) professional development activities that you may have participated in beginning [REFERENCE DATE] (such as workshops, institutes, webinars, mentoring/coaching and leadership activities).

A1. <u>Since [REFERENCE DATE]</u>, did you participate in any professional development (PD) activities related to English/Language Arts (ELA)?

Professional development is defined as in-service training to increase content knowledge and instructional skills of teachers. Some methods of PD include workshops (in person or web-based), one-on-one coaching, team teaching, college courses, teacher observation and feedback.

- ❍ Yes 1 A2
- ❍ No 2 A5
- ❍ Don't Know ..... D A5

NO RESPONSE ..... M A2

A2. <u>Since [REFERENCE DATE]</u>, which ELA professional development activities were offered to teachers at your school?

Select one response per row

		OFFERED	NOT OFFERED	DON'T KNOW
a.	Support for college course taken after your first certification	1 ☐	0 ☐	d ☐
b.	Workshop or training session	1 ☐	0 ☐	d ☐
c.	Conference or professional association meeting	1 ☐	0 ☐	d ☐
d.	Committee or task force focusing on curriculum, instruction, or student assessment	1 ☐	0 ☐	d ☐
e.	Regularly scheduled discussion group or study group	1 ☐	0 ☐	d ☐
f.	Teacher collaborative or network, such as one organized by an outside agency or over the Internet	1 ☐	0 ☐	d ☐
g.	Individual or collaborative research	1 ☐	0 ☐	d ☐
h.	Opportunities for reading about curriculum, instruction or assessment on a regular basis—for example, reading educational journals, books, or the Internet at least once a month	1 ☐	0 ☐	d ☐
i.	Co-teaching/team teaching with a lead/master/cooperating teacher	1 ☐	0 ☐	d ☐
j.	Consultation with a ELA specialist/instructional coach*	1 ☐	0 ☐	d ☐
k.	Grade level or subject area meetings/discussion groups	1 ☐	0 ☐	d ☐

*ELA specialists/instructional coaches assist in developing and implementing instructional plans that align with district goals and curriculum; conduct teacher observations and provide feedback; and also may provide in-person and/or online individual and/or group instructional coaching and mentoring to teachers to improve classroom instruction.

A2L Since [REFERENCE DATE], were any other ELA PD activities offered at your school?

☐ Yes 1 A2lspec

☐ No 0 A3

☐ Don't Know ..... D A3

NO RESPONSE ..... M A3

A2lspec. What other ELA PD activities were offered at your school?

GO TO A3

A3. Since [REFERENCE DATE], have you participated in any of the following ELA professional development activities?

Select one response per row

	PARTICIPATED	DID NOT PARTICIPATE	DON'T KNOW
a. College course taken after your first certification	1 ☐	0 ☐	d ☐
b. Workshop or training session	1 ☐	0 ☐	d ☐
c. Conference or professional association meeting	1 ☐	0 ☐	d ☐
d. Committee or task force focusing on curriculum, instruction or student assessment	1 ☐	0 ☐	d ☐
e. Regularly scheduled discussion group or study group	1 ☐	0 ☐	d ☐
f. Teacher collaborative or network, such as one organized by an outside agency or over the Internet	1 ☐	0 ☐	d ☐
g. Individual or collaborative research	1 ☐	0 ☐	d ☐
h. Reading about curriculum, instruction or assessment on a regular basis—for example, reading educational journals, books, or the Internet at least once a month	1 ☐	0 ☐	d ☐
i. Co-teaching/team teaching with a lead/master/cooperating teacher	1 ☐	0 ☐	d ☐
j. Consultation with a ELA specialist/instructional coach*	1 ☐	0 ☐	d ☐
k. Grade level or subject area meetings/discussion groups	1 ☐	0 ☐	d ☐

* ELA specialists/instructional coaches assist in developing and implementing instructional plans that align with district goals and curriculum; conduct teacher observations and provide feedback; also may provide in-person and/or online individual and/or group instructional coaching and mentoring to ELA teachers to improve classroom instruction.

A3L. Since [REFERENCE DATE], did you participate in any other ELA PD activities offered at your school?

- ❍ Yes 1 A3lspec
- ❍ No 0 A4
- ❍ Don't Know ..... D A4

NO RESPONSE ..... M A4

A3lspec. In what other ELA PD activities did you participate at your school?

GO TO A4

The next set of questions ask about topics that you may have encountered when you participated in professional development activities since [REFERENCE DATE].

Professional development can be defined as training to increase content knowledge and instructional skills of teachers. Some methods of PD include workshops (in person or web-based), one-on-one coaching, team teaching, college courses, teacher observation and feedback.

A4a1. Since [REFERENCE DATE], did you participate in professional development that covered the topic of "implementing a Common Core-aligned curriculum?"

- ❍ Yes ..... 1 GO TO A4b1
- ❍ No ..... 0 GO TO A4a2

NO RESPONSE ..... M GO TO A4a2

A4b1. How useful to your teaching was the professional development that focused on the topic "implementing a Common Core-aligned curriculum?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

NO RESPONSE ..... M

A4a2. Since [REFERENCE DATE], did you participate in professional development that covered the topic of "engaging students in discussion and thinking protocols?"

❍ Yes	1	GO TO A4b2
❍ No	0	GO TO A4a3
NO RESPONSE	M	GO TO A4a3

A4b2. How useful to your teaching was the professional development that focused on the topic "engaging students in discussion and thinking protocols?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

A4a3. Since [REFERENCE DATE], did you participate in professional development that covered the topic of "helping students read complex text through close reading and other strategies?"

❍ Yes	1	GO TO A4b3
❍ No	0	GO TO A4a4
NO RESPONSE	M	GO TO A4a4

A4b3. How useful to your teaching was the professional development that focused on the topic "helping students read complex text through close reading and other strategies?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

A4a4. Since [REFERENCE DATE], did you participate in professional development that covered the topic of "helping students select and provide evidence for their thinking?"

- ❍ Yes ..... 1 GO TO A4b4
- ❍ No ..... 0 GO TO A4a5

NO RESPONSE ..... M GO TO A4a5

A4b4. How useful to your teaching was the professional development that focused on the topic "helping students select and provide evidence for their thinking?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

A4a5. Since [REFERENCE DATE], did you participate in professional development that covered the topic of "helping students learn vocabulary?"

- ❍ Yes ..... 1 GO TO A4b5
- ❍ No ..... 0 GO TO A4a6

NO RESPONSE ..... M GO TO A4a6

A4b5. How useful to your teaching was the professional development that focused on the topic "helping students learn vocabulary?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

A4a6. Since [REFERENCE DATE], did you participate in professional development that covered the topic of "engaging students in the analysis of models of writing?"

- ❍ Yes ..... 1 GO TO A4b6
- ❍ No ..... 0 GO TO A4a7

NO RESPONSE ..... M GO TO A4a7

A4b6. How useful to your teaching was the professional development that focused on the topic "engaging students in the analysis of models of writing?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

A4a7. Since [REFERENCE DATE], did you participate in professional development that covered the topic of "using curriculum-based and other data to drive instruction (student's writing and assessments)?"

- ❍ Yes ..... 1 GO TO A4b7
- ❍ No ..... 0 GO TO A4a8

NO RESPONSE ..... M GO TO A4a8

A4b7. How useful to your teaching was the professional development that focused on the topic "using curriculum-based and other data to drive instruction (student's writing and assessments)?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

A4a8. <u>Since [REFERENCE DATE],</u> did you participate in professional development that covered the topic of "developing writing from sources?"

- ❍ Yes ..... 1 GO TO A4b8
- ❍ No ..... 0 GO TO A4a9

NO RESPONSE ..... M GO TO A4a9

A4b8. How useful to your teaching was the professional development that focused on the topic "<u>developing writing from sources?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

A4a9. <u>Since [REFERENCE DATE],</u> did you participate in professional development that covered the topic of "supporting English Language Learners?"

- ❍ Yes ..... 1 GO TO A4b9
- ❍ No ..... 0 GO TO A4a10

NO RESPONSE ..... M GO TO A4a10

A4b9. How useful to your teaching was the professional development that focused on the topic "<u>supporting English Language Learners?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

A4a10. <u>Since [REFERENCE DATE],</u> did you participate in professional development that covered the topic of "supporting Special Education Students?"

❍ Yes	1	GO TO A4b10
❍ No	0	GO TO A4a11
NO RESPONSE	M	GO TO A4a11

A4b10. How useful to your teaching was the professional development that focused on the topic "<u>supporting Special Education Students?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

A4a11. <u>Since [REFERENCE DATE],</u> did you participate in professional development that covered the topic of "using learning targets?"

❍ Yes	1	GO TO A4b11
❍ No	0	GO TO A4a12
NO RESPONSE	M	GO TO A4a12

A4b11. How useful to your teaching was the professional development that focused on the topic "<u>using learning targets?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

A4a12. <u>Since [REFERENCE DATE],</u> did you participate in professional development that covered the topic of "using formative assessment strategies?"

- ❍ Yes ..... 1 GO TO A4b12
- ❍ No ..... 0 GO TO A4a13

NO RESPONSE ..... M GO TO A4a13

A4b12. How useful to your teaching was the professional development that focused on the topic "<u>using formative assessment strategies?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

A4a13. <u>Since [REFERENCE DATE],</u> did you participate in professional development that covered the topic of "supporting students in tracking their learning and setting goals?"

- ❍ Yes ..... 1 GO TO A4b13
- ❍ No ..... 0 GO TO A4a14

NO RESPONSE ..... M GO TO A4a14

A4b13. How useful to your teaching was the professional development that focused on the topic "<u>supporting students in tracking their learning and setting goals?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

A4a14. <u>Since [REFERENCE DATE],</u> did you participate in professional development that covered the topic of "providing peer critique and other feedback?"

❍ Yes	1	GO TO A4b14
❍ No	0	GO TO A4a15
NO RESPONSE	M	GO TO A4a15

A4b14. How useful to your teaching was the professional development that focused on the topic "<u>providing peer critique and other feedback?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

A4a15. <u>Since [REFERENCE DATE],</u> did you participate in professional development that covered the topic of "co-constructing criteria for success (academic and behavioral)?"

❍ Yes	1	GO TO A4b15
❍ No	0	GO TO A4a16
NO RESPONSE	M	GO TO A4a16

A4b15. How useful to your teaching was the professional development that focused on the topic "<u>co-constructing criteria for success (academic and behavioral)?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

A4a16. Since [REFERENCE DATE], did you participate in professional development that covered the topic of "managing an active classroom where students are active participants/leaders?"

- ❍ Yes ..... 1 GO TO A4b16
- ❍ No ..... 0 GO TO A5

NO RESPONSE ..... M GO TO A5

A4b16. How useful to your teaching was the professional development that focused on the topic "managing an active classroom where students are active participants/leaders?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

A5. For the list below, please indicate whether any of the following coaching or mentoring activities related to teaching ELA were offered at your school from [REFERENCE DATE] until now.

Select one response per row

	OFFERED	NOT OFFERED	DON'T KNOW
a. Opportunity to observe a teacher in my school.	1 ☐	0 ☐	d ☐
b. Opportunity to observe a teacher in another school.	1 ☐	0 ☐	d ☐
c. Opportunity to be observed by a teacher from my school.	1 ☐	0 ☐	d ☐
d. Opportunity to provide mentoring and/or coaching to another teacher in my school.	1 ☐	0 ☐	d ☐
e. Opportunity to receive mentoring and/or coaching from another teacher in my school.	1 ☐	0 ☐	d ☐
f. Opportunity to receive online mentoring and/or coaching from an instructional coach (i.e., Video-based coaching).	1 ☐	0 ☐	d ☐
g. Opportunity to receive in-person mentoring and/or coaching from an instructional coach.	1 ☐	0 ☐	d ☐

A6. Since [REFERENCE DATE], have you participated in any of the following coaching or mentoring activities related to teaching ELA?

Select one response per row

	PARTICIPATED	DID NOT PARTICIPATE	DON'T KNOW
a. I observed a teacher in my school.	1 ☐	0 ☐	d ☐
b. I observed a teacher in another school.	1 ☐	0 ☐	d ☐
c. I was observed by a teacher from my school.	1 ☐	0 ☐	d ☐
d. I was observed by a teacher from another school	1 ☐	0 ☐	d ☐
e. I provided mentoring and/or coaching to another teacher in my school.	1 ☐	0 ☐	d ☐
f. I received mentoring and/or coaching from another teacher in my school.	1 ☐	0 ☐	d ☐
g. I received online mentoring and/or coaching from an instructional coach (i.e., Video-based coaching.)	1 ☐	0 ☐	d ☐
h. I received in-person mentoring and/or coaching from an instructional coach.	1 ☐	0 ☐	d ☐

A7. For each of the special leadership roles listed below, please indicate if you have any of these special leadership responsibilities for ELA at your school.

Leadership responsibilities include being a mentor teacher, lead teacher, resource specialist, departmental chair, or master teacher.

Select one response per row

	SELECT ONE RESPONSE PER ROW			
	YES	NO	NOT OFFERED AT MY SCHOOL	DON'T KNOW
a. Coaching or mentoring one or more peers/colleagues	1 ☐	0 ☐	n ☐	d ☐
b. Grade-level team leader (provides leadership and support to teachers within their grade level to ensure the team meets student achievement goals; acts as a liaison between school's leadership and teachers in their grade level; leads and coordinates grade-level team meetings, organizational practices, parent communication, and grade-wide activities; may also provide instructional support to teachers in the grade level)	1 ☐	0 ☐	n ☐	d ☐
c. Department chair (manages the department and provides instructional leadership, ensures curriculum alignment with state standards, leads in the development and implementation of standard assessments)	1 ☐	0 ☐	n ☐	d ☐
d. Resource specialist (provides instruction and services to students who have an individualized educational program and who are assigned to regular classroom teachers for the majority of a school day)	1 ☐	0 ☐	n ☐	d ☐
e. Master teacher (demonstrates excellence inside and outside of the classroom through consistent leadership and focused collaboration with students, parents, other teachers and administration to maximize student learning; strives for distinguished teaching through continued professional growth)	1 ☐	0 ☐	n ☐	d ☐
f. Instructional specialist (helps teachers implement effective teaching strategies)	1 ☐	0 ☐	n ☐	d ☐
g. Curriculum specialist (helps teachers understand content standards, implement the current curriculum, and develop, understand and implement shared student assessments)	1 ☐	0 ☐	n ☐	d ☐
h. Data specialist (helps teachers understand, analyze, and use data to help improve instruction)	1 ☐	0 ☐	n ☐	d ☐
i. Other *(specify)* ☐	1 ☐	0 ☐	n ☐	d ☐

B. OVERALL INSTRUCTIONA L PRACTICES

This section collects information about: planning instruction, selecting instructional materials/texts, development of student activities, tasks and student assessments, instructional techniques/approaches and student behavior/classroom management.

B1. When planning instruction, which source documents do you use?

SELECT ALL THAT APPLY

1 ☐ School department goals
2 ☐ EL Education curriculum
3 ☐ District curriculum
4 ☐ State standards
5 ☐ Common Core State Standards
6 ☐ Student identified documents
7 ☐ Information from assessments
8 ☐ Other curriculum materials
9 ☐ Other source documents not listed above

IF YOUR SCHOOL IS PARTICIPATING IN EL TPP, GO TO B6

B2. This year, how much control do you have in your classroom over selecting texts and other instructional materials?

Select one only

1 ☐ Full control: I decide what texts and materials to use.
2 ☐ A good deal of control: I select from among resources available in my school and/or district what texts and materials to use.
3 ☐ Some control: The main texts and materials are selected by administration, but I can use supplemental texts.
4 ☐ No control: All texts and materials are selected by administration → **GO TO B4 ON PAGE 8**

B3. When you <u>select texts and other instructional materials</u>, how important are the following factors?

Select one response per row

		NOT AT ALL IMPORTANT	SOMEWHAT IMPORTANT	IMPORTANT	VERY IMPORTANT
a.	Students' reading levels (based on proficiency test scores)	1 ☐	2 ☐	3 ☐	4 ☐
b.	Students' actual grade levels	1 ☐	2 ☐	3 ☐	4 ☐
c.	Type of text (literary, such as poetry, plays, stories, novels; informational such as textbooks, newspaper articles, essays, biographies)	1 ☐	2 ☐	3 ☐	4 ☐
d.	Complexity of text (difficulty of content/concepts, vocabulary, length, syntax)	1 ☐	2 ☐	3 ☐	4 ☐
e.	Students' interests	1 ☐	2 ☐	3 ☐	4 ☐
f.	District/school curriculum requirements	1 ☐	2 ☐	3 ☐	4 ☐
g.	Increasing the complexity of texts across the school year				
h.	ELA Department recommendations	1 ☐	2 ☐	3 ☐	4 ☐
i.	Availability of materials	1 ☐	2 ☐	3 ☐	4 ☐
j.	State standards	1 ☐	2 ☐	3 ☐	4 ☐
k.	Common Core State Standards	1 ☐	2 ☐	3 ☐	4 ☐
l.	Students' special needs (need for large print or audio versions of texts/materials)	1 ☐	2 ☐	3 ☐	4 ☐

B4. When you develop activities or tasks for students to engage in, how important are the following factors?

Select one response per row

	NOT AT ALL IMPORTANT	SOMEWHAT IMPORTANT	IMPORTANT	VERY IMPORTANT
a. Students' learning styles	1 ☐	2 ☐	3 ☐	4 ☐
b. Working with students to set learning targets	1 ☐	2 ☐	3 ☐	4 ☐
c. Providing opportunities for students' self-assessments	1 ☐	2 ☐	3 ☐	4 ☐
d. Differentiating instruction	1 ☐	2 ☐	3 ☐	4 ☐
e. Ensuring students refer to text when answering questions	1 ☐	2 ☐	3 ☐	4 ☐
f. Providing learning that has meaning beyond the classroom	1 ☐	2 ☐	3 ☐	4 ☐
g. Providing learning that is a process of investigation and discovery	1 ☐	2 ☐	3 ☐	4 ☐
h. Providing opportunities for students to collaborate	1 ☐	2 ☐	3 ☐	4 ☐
i. Providing opportunities for students to write from sources	1 ☐	2 ☐	3 ☐	4 ☐
j. Providing learning that is a collaborative process between teachers, students and other organizations (e.g. community organizations, academic organizations, online communities)	1 ☐	2 ☐	3 ☐	4 ☐
k. Providing students with opportunities to present or perform their work	1 ☐	2 ☐	3 ☐	4 ☐
l. Providing regular practice with increasingly complex texts and academic language	1 ☐	2 ☐	3 ☐	4 ☐
m. Providing opportunities for students to improve vocabulary through conversation, direct instruction, and reading	1 ☐	2 ☐	3 ☐	4 ☐
n. Integrating technology into the classroom	1 ☐	2 ☐	3 ☐	4 ☐

B5. When you develop student assessments, how important are the following factors?

Select one response per row

	NOT AT ALL IMPORTANT	SOMEWHAT IMPORTANT	IMPORTANT	VERY IMPORTANT
a. Getting students involved in developing questions and/or rubrics	1 ☐	2 ☐	3 ☐	4 ☐
b. Ensuring questions require students to refer to text	1 ☐	2 ☐	3 ☐	4 ☐
c. Ensuring selected texts are challenging	1 ☐	2 ☐	3 ☐	4 ☐
d. Ensuring higher order thinking	1 ☐	2 ☐	3 ☐	4 ☐
e. Ensuring student creativity	1 ☐	2 ☐	3 ☐	4 ☐
f. Ensuring opportunity for student self-expression	1 ☐	2 ☐	3 ☐	4 ☐

B5g Are there any other factor(s) you think are important when you develop student assessments?

- ❍ Yes 1 B5gspec
- ❍ No 0 B6
- ❍ Don't Know ..... D B6

NO RESPONSE ..... M B6

B5gspec What other factor(s) do you think are important?

GO TO B5grate

B5grate How important do you think these factor(s) are?

Select one only

- ❍ Not at all important 1
- ❍ Somewhat important 2
- ❍ Important 3
- ❍ Very important 4

CONTINUE TO B6

B6. Below are some common approaches to teaching ELA. Please select the one that best describes <u>your</u> own approach to instruction.

Var

Select one only

- ❍ I teach particular books, short stories, essays, and poems that I think students should read and then I organize instruction around them, teaching a variety of reading skills and strategies as tools for students to understand the texts. ..... 1 GO TO B7
- ❍ I focus instruction on reading skills and strategies first, and then have students apply these skills and strategies to any book, short story, essay, or poem they read. ..... 2 GO TO B7
- ❍ I focus instruction on topics that build students' content knowledge, selecting texts about specific topics, and standards that can be taught through those texts. ..... 3 GO TO B7
- ❍ I focus instruction on developing students reading and writing skills as well as developing their abilities to engage in higher-order thinking. ..... GO TO B7
- ❍ Other ..... 4 GO TO B6a
- ❍ Not sure ..... 5 GO TO B7

NO RESPONSE ..... M GO TO B7

B6a. What other approaches to instruction have you used to teach ELA?

[ ]

NO RESPONSE ..... M

B7. <u>In the past two weeks of regular ELA instruction</u>, how often did you do the following?

Select one response per row

	NEVER	ONCE OR TWICE A WEEK	ALMOST EVERY DAY	EVERYDAY	MORE THAN ONCE A DAY
a. Divided students into groups or teams by readiness or ability levels	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
b. Set different achievement standards or targets for some students	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
c. Used various teaching methods to meet students' individual needs	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
d. Supplemented the regular course curriculum with additional materials for some students	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
e. Provided some students with different texts	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
f. Provided some students with additional scaffolding or support	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
g. Paced instruction differently for some students	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

B8. <u>In the past two weeks of regular ELA instruction</u>, how often did the following occur?

Select one response per row

	NEVER	ONCE OR TWICE A WEEK	ALMOST EVERY DAY	EVERYDAY	MORE THAN ONCE A DAY
a. Students provided feedback on each other's work.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
b. Students worked with other students.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
c. Students asked for more challenging work.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
d. Students checked their own progress against learning targets.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
e. Students participated in developing rubrics or evaluation criteria	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
f. Students gave input in setting the learning targets.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
g. Students shared their work with their peers.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

B9. How much do you agree or disagree with the following statements about <u>your</u> ELA class(es)?

Select one response per row

	DISAGREE	SOMEWHAT DISAGREE	SOMEWHAT AGREE	AGREE
a. I set clear expectations for student behavior.	1 ☐	2 ☐	3 ☐	4 ☐
b. I require that students in my classroom follow the rules at all times.	1 ☐	2 ☐	3 ☐	4 ☐
c. I redirect students back to the topic when they get off-task.	1 ☐	2 ☐	3 ☐	4 ☐
e. I manage my class very well.	1 ☐	2 ☐	3 ☐	4 ☐
f. I often acknowledge positive student behavior.	1 ☐	2 ☐	3 ☐	4 ☐
g. I often emphasize positive character traits and values in the classroom.	1 ☐	2 ☐	3 ☐	4 ☐
h. I encourage a collaborative classroom environment.	1 ☐	2 ☐	3 ☐	4 ☐

B10. <u>During a typical class period</u>, how often do the following events take time away from instruction?

Select one response per row

	NEVER	ONCE OR TWICE A WEEK	ALMOST EVERY DAY	EVERYDAY	MORE THAN ONCE A DAY
a. An individual student or students disrupts class for less than 5 minutes.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
b. An individual student or students disrupts class for 5 to 10 minutes.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
c. An individual student or students disrupts class for more than 10 minutes.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
d. Transitions from one activity to another take more than 5 minutes.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
e. Announcements on the PA system disrupt class for more than 2 minutes.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
f. Noise from outside the classroom disrupts class for more than 2 minutes.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

C. READING INSTRUCTIONAL PRACTICES

For the next set of questions, please indicate if you have used the following types of texts with your students during the past two weeks of regular instruction and if you have, please enter the number of texts used in your classroom and the name of the text you most recently used.

If you teach more than one grade level of ELA, please think of one grade level when responding to this question.

C1a1. During the past two weeks of regular instruction, have you used 'novels' in your classroom?

- ❍ Yes ..... 1 GO TO C1b1
- ❍ No ..... 0 GO TO C1a2

NO RESPONSE ..... M GO TO C1a2

C1b1. How many novels have you used in the past two weeks of regular instruction?

[ ] GO TO C1c1

NO RESPONSE ..... M GO TO C1c1

C1c1. What is the name of the novel most recently used in your classroom?

[ ]

NO RESPONSE ..... M

C1a2. During the past two weeks of regular instruction, have you used 'novellas' in your classroom?

- ❍ Yes ..... 1 GO TO C1b2
- ❍ No ..... 0 GO TO C1a3

NO RESPONSE ..... M GO TO C1a3

C1b2. How many novellas have you used in the past two weeks of regular instruction?

[ ] GO TO C1c1

NO RESPONSE ..... M GO TO C1c1

C1c2. What is the name of the novella most recently used in your classroom?

[ ]

NO RESPONSE ..... M

C1a3. During the past two weeks of regular instruction, have you used 'fictional short stories' in your classroom?

- ❍ Yes ..... 1 GO TO C1b3
- ❍ No ..... 0 GO TO C1a4

NO RESPONSE ..... M GO TO C1a4

C1b3. How many fictional short stories have you used in the past two weeks of regular instruction?

[ ] GO TO C1c3

NO RESPONSE ..... M GO TO C1c3

C1c3. What is the name of the fictional short story most recently used in your classroom?

[ ]

NO RESPONSE ..... M

C1a4. During the past two weeks of regular instruction, have you used 'poetry in your classroom?

- ❍ Yes ..... 1 GO TO C1b4
- ❍ No ..... 0 GO TO C1a5

NO RESPONSE ..... M GO TO C1a5

C1b4. How many works of poetry have you used in the past two weeks of regular instruction?

[ ] GO TO C1c4

NO RESPONSE ..... M GO TO C1c4

C1c4. What is the name of the work of poetry most recently used in your classroom?

[ ]

NO RESPONSE ..... M

C1a5. During the past two weeks of regular instruction, have you used 'dramas/plays' in your classroom?

- ❍ Yes ..... 1 GO TO C1b5
- ❍ No ..... 0 GO TO C1a6

NO RESPONSE ..... M GO TO C1a6

C1b5. How many dramas/plays have you used in the past two weeks of regular instruction?

[ ] GO TO C1c5

NO RESPONSE ..... M GO TO C1c5

C1c5. What is the name of the drama/play most recently used in your classroom?

[ ]

NO RESPONSE ..... M

C1a6. During the past two weeks of regular instruction, have you used 'biographies/autobiographies' in your classroom?

- ❍ Yes ..... 1 GO TO C1b6
- ❍ No ..... 0 GO TO C1a7

NO RESPONSE ..... M GO TO C1a7

C1b6. How many biographies/autobiographies have you used in the past two weeks of regular instruction?

[ ] GO TO C1c6

NO RESPONSE ..... M GO TO C1c6

C1c6. What is the name of the biography/autobiography most recently used in your classroom?

[ ]

NO RESPONSE ..... M

C1a7. <u>During the past two weeks of regular instruction</u>, have you used 'personal essays' in your classroom?

- ❍ Yes ..... 1 GO TO C1b7
- ❍ No ..... 0 GO TO C1a8

NO RESPONSE ..... M GO TO C1a8

C1b7. How many <u>personal essays</u> have you used in the past two weeks of regular instruction?

[ ] GO TO C1c7

NO RESPONSE ..... M GO TO C1c7

C1c7. What is the name of the personal essay most recently used in your classroom?

[ ]

NO RESPONSE ..... M

C1a8. <u>During the past two weeks of regular instruction</u>, have you used 'memoirs/diaries' in your classroom?

- ❍ Yes ..... 1 GO TO C1b8
- ❍ No ..... 0 GO TO C1a9

NO RESPONSE ..... M GO TO C1a9

C1b8. How many <u>memoirs/diaries</u> have you used in the past two weeks of regular instruction?

[ ] GO TO C1c8

NO RESPONSE ..... M GO TO C1c8

C1c8. What is the name of the memoir/diary most recently used in your classroom?

[ ]

NO RESPONSE ..... M

C1a9. <u>During the past two weeks of regular instruction</u>, have you used 'nonfictional short stories' in your classroom?

- ❍ Yes ..... 1 GO TO C1b9
- ❍ No ..... 0 GO TO C1a10

NO RESPONSE ..... M GO TO C1a10

C1b9. How many <u>nonfictional short stories</u> have you used in the past two weeks of regular instruction?

[ ] GO TO C1c9

NO RESPONSE ..... M GO TO C1c9

C1c9. What is the name of the nonfictional short story most recently used in your classroom?

[ ]

NO RESPONSE ..... M

C1a10. <u>During the past two weeks of regular instruction</u>, have you used 'speeches' in your classroom?

- ❍ Yes ..... 1 GO TO C1b10
- ❍ No ..... 0 GO TO C1a11

NO RESPONSE ..... M GO TO C1a11

C1b10. How many <u>speeches</u> have you used in the past two weeks of regular instruction?

[ ] GO TO C1c10

NO RESPONSE ..... M GO TO C1c10

C1c10. What is the name of the speech most recently used in your classroom?

[ ]

NO RESPONSE ..... M

C1a11. <u>During the past two weeks of regular instruction</u>, have you used articles from magazines, newspapers, journals, or the internet in your classroom?

- ❍ Yes ..... 1 GO TO C1b11
- ❍ No ..... 0 GO TO C1a12

NO RESPONSE ..... M GO TO C1a12

C1b11. <u>During the past two weeks of regular instruction</u>, how many articles have you used from magazines, newspapers, journals or the internet in the past two weeks of regular instruction?

[ ] GO TO C1c11

NO RESPONSE ..... M GO TO C1c11

C1c11. What is the name of the magazine, newspaper, journal, or internet article most recently used in your classroom?

[ ]

NO RESPONSE ..... M

C1a12. <u>During the past two weeks of regular instruction</u>, have you used 'textbooks/technical writings' in your classroom?

- ❍ Yes ..... 1 GO TO C1b12
- ❍ No ..... 0 GO TO C1a13

NO RESPONSE ..... M GO TO C1a13

C1b12. How many <u>textbooks/technical writings</u> have you used in the past two weeks of regular instruction?

[ ] GO TO C1c12

NO RESPONSE ..... M GO TO C1c12

C1c12. What is the name of the textbook/technical writings most recently used in your classroom?

[ ]

NO RESPONSE ..... M

C1a13. <u>During the past two weeks of regular instruction</u>, have you used 'reference books (dictionary, thesaurus, encyclopedia, atlas)' in your classroom?

- ❍ Yes ..... 1 GO TO C1b13
- ❍ No ..... 0 GO TO C1a14

NO RESPONSE ..... M GO TO C1a14

C1b13. How many <u>reference books (dictionary, thesaurus, encyclopedia, atlas)</u> have you used in the past two weeks of regular instruction?

[ ] GO TO C1c13

NO RESPONSE ..... M GO TO C1c13

C1c13. What is the name of the reference book (dictionary, thesaurus, encyclopedia, atlas) most recently used in your classroom?

[ ]

NO RESPONSE ..... M

C1a14. <u>During the past two weeks of regular instruction</u>, have you used any other types of texts in your classroom?

- ❍ Yes (Please specify) ..... 1 GO TO C1b14

Specify [ ]

- ❍ No ..... 2 GO TO C2

NO RESPONSE ..... M GO TO C2

C1b14. How many <u>other types of text</u> have you used in the past two weeks of regular instruction?

[ ] GO TO C1c14

NO RESPONSE ..... M GO TO C1c14

C1c14. What is the name of the other type of text most recently used in your classroom?

[ ]

NO RESPONSE ..... M

C2. How many of your students are struggling readers (reads 3 or more years below grade level)?

Select one only

- ❍ None ..... 1 GO TO C3
- ❍ About one quarter or less 2
- ❍ About one quarter to half 3
- ❍ More than half 4
- ❍ Almost all 5

C2a. For the struggling readers in your class(es), do you use abridged or adapted versions of texts or do all students read the same version of the texts?

SELECT ONE RESPONSE PER ROW

	NEVER OR HARDLY EVER	SOMETIMES	OFTEN	VERY OFTEN OR ALWAYS
a. I use the same unabridged or un-adapted version of a text for all students.	1 ☐	2 ☐	3 ☐	4 ☐
b I use abridged or adapted versions for all students.	1 ☐	2 ☐	3 ☐	4 ☐
c I use abridged or adapted versions for struggling readers.	1 ☐	2 ☐	3 ☐	4 ☐
d. I adapt or modify text myself for struggling readers.	1 ☐	2 ☐	3 ☐	4 ☐

C3. How many texts do you usually use per ELA lesson?

|___|___| NUMBER OF TEXTS

C4. <u>During the past two weeks</u>, about how many pages of text did you have your students <u>read in class</u>?

Select one only

- ❍ 1 to 5 pages 1
- ❍ 6 to 10 oages 2
- ❍ 11 to 15 pages 3
- ❍ 16 to 20 pages 4
- ❍ More than 20 pages 5

C5. <u>During the past two weeks</u>, about how much time did your students spend reading independently in class (total time spent across the past ten days of regular instruction)?

Select one only

- ❍ None 1
- ❍ Less than a half hour 2
- ❍ Between half an hour and an hour 3
- ❍ One to two hours 4
- ❍ More than 2 hours 5

C6. **During the past two weeks, about how many pages of text did you have your students read when not in school (total time spent across the past two weeks)?**

Select one only

- ❍ 1 to 5 pages 1
- ❍ 6 to 10 oages 2
- ❍ 11 to 15 pages 3
- ❍ 16 to 20 pages 4
- ❍ More than 20 pages 5

C7. **During the past two weeks, about how much time, in total, did your students spend reading independently when not in school (total time spent across the past two weeks)?**

Select one only

- ❍ None 1
- ❍ Less than a half hour 2
- ❍ Between half an hour and an hour 3
- ❍ One to two hours 4
- ❍ More than 2 hours 5
- ❍ Don't know ..... d

C8. **During the past two weeks, how often did students do the following tasks (total time spent across the past ten days of regular instruction)??**

SELECT ONE RESPONSE PER ROW

	NEVER	ONCE OR TWICE A WEEK	ALMOST EVERY DAY	EVERYDAY	MORE THAN ONCE A DAY
a. Read required materials	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
b. Read materials of their own choosing	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
c. Discuss texts they have read with partners or a small group of students	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
d. Write about texts they have read	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
e. Produce extended writing by incorporating key details from texts they have read	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
f. Share their ideas about and/or understanding of the texts they read with the whole class	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

C9. How often do you focus on the following processes when students read texts in class?

SELECT ONE RESPONSE PER ROW

	NEVER OR HARDLY EVER	SOMETIMES	OFTEN	VERY OFTEN OR ALWAYS
a. Determine central idea(s) or theme(s) of the text	1 ☐	2 ☐	3 ☐	4 ☐
b. Determine the author's point of view or purpose	1 ☐	2 ☐	3 ☐	4 ☐
c. Cite textual evidence to support analysis of the text	1 ☐	2 ☐	3 ☐	4 ☐
d. Critique or evaluate arguments and specific claims in context	1 ☐	2 ☐	3 ☐	4 ☐
e. Analyze individuals, events and ideas in the text (e.g., how they are introduced, how they interact with each other, how text makes connections/distinctions among them)	1 ☐	2 ☐	3 ☐	4 ☐
f. Summarize the text	1 ☐	2 ☐	3 ☐	4 ☐
g. Determine meaning of words and phrases in context	1 ☐	2 ☐	3 ☐	4 ☐
h. Analyze the structure of sentences, paragraphs and major sections of texts	1 ☐	2 ☐	3 ☐	4 ☐
i. Locate or recall information to answer questions to demonstrate understanding of the text	1 ☐	2 ☐	3 ☐	4 ☐
j. Integrate or interpret information across sections of the text	1 ☐	2 ☐	3 ☐	4 ☐
k. Integrate or compare and contrast information in different media or formats	1 ☐	2 ☐	3 ☐	4 ☐
l. Analyze differences and similarities in perspectives between multiple texts on the same topic	1 ☐	2 ☐	3 ☐	4 ☐
m. Relate the work, its characters, and/or its themes to their own lives	1 ☐	2 ☐	3 ☐	4 ☐
n. Relate the story or literary work, its characters, and/or its themes to something they have read before	1 ☐	2 ☐	3 ☐	4 ☐

C10. <u>During a typical week</u>, how often do you engage in the following activities?

SELECT ONE RESPONSE PER ROW

	NEVER	ONCE OR TWICE A WEEK	ALMOST EVERY DAY	EVERYDAY	MORE THAN ONCE A DAY
a. Ask students to read aloud	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
b. Ask students to read silently	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
c. Ask students to read in pairs or in small groups	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
d. Give students time to read books they have chosen themselves	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
e. Ask students to talk with each other about what they have read	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
f. Ask students to write about something they have read	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
g. Ask students to work in a reading notebook or on a skills worksheet, note catcher or graphic organizer	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
h. Ask students to do a group activity or project about what they have read	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
i. Ask students to discuss different interpretations of what they have read	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
j. Ask students to explain or support their understanding of what they have read	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
k. Watch movies, videos, television shows; or listen to music	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
l. Help students understand new words	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
m. Ask students to make predictions about what they read as they are reading	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
n. Ask students to make generalizations and draw inferences based on what they have read	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
o. Ask students to describe the style or structure of the text they have read	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
p. Ask students to grapple with difficult text on their own	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

C11. <u>In the past two weeks</u>, how often did you do each of the following?

SELECT ONE RESPONSE PER ROW

	NEVER	ONCE OR TWICE A WEEK	ALMOST EVERY DAY	EVERYDAY	MORE THAN ONCE A DAY
a. Provided direct instruction in overall reading strategies and/or strategies specific to reading fiction or non-fiction	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
b. Modeled for students overall reading strategies and/or strategies specific to reading fiction or non-fiction	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
c. Provided opportunities for students to practice overall reading strategies and/or strategies specific to reading fiction or non-fiction	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
d. Provided feedback to students on their use of overall reading strategies and/or strategies specific to reading fiction or non-fiction	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
e. Asked students to relate what they read to their own experience or to something they have learned before	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

D. WRITING INSTRUCTIONAL PRACTICES

This section collects information about your writing instruction practices such as: grading/evaluating student writing, instructional techniques for writing, writing purposes and writing assignments.

D1. When you provide students with feedback on their writing, how important are the following?

SELECT ONE RESPONSE PER ROW

	NOT AT ALL IMPORTANT	SOMEWHAT IMPORTANT	IMPORTANT	VERY IMPORTANT
a. Development of ideas and use of evidence to support ideas	1 ☐	2 ☐	3 ☐	4 ☐
b. Organization of ideas	1 ☐	2 ☐	3 ☐	4 ☐
c. Effective use of language (e.g., sentence variety, word choice, tone)	1 ☐	2 ☐	3 ☐	4 ☐
d. Mechanics (e.g., spelling, capitalization, punctuation, grammar) and conventions, paragraphing, using appropriate formats, such as subheadings and quotations	1 ☐	2 ☐	3 ☐	4 ☐
e. For text-based writing, the accuracy or appropriateness of their references to texts	1 ☐	2 ☐	3 ☐	4 ☐
f. Quality of thinking	1 ☐	2 ☐	3 ☐	4 ☐
g. Demonstration of content knowledge or understanding of text	1 ☐	2 ☐	3 ☐	4 ☐

D2. When grading or evaluating your students' writing, how important is each of the following?

SELECT ONE RESPONSE PER ROW

	NOT AT ALL IMPORTANT	SOMEWHAT IMPORTANT	IMPORTANT	VERY IMPORTANT
a. Development of ideas and use of evidence to support ideas	1 ☐	2 ☐	3 ☐	4 ☐
b. Organization of ideas	1 ☐	2 ☐	3 ☐	4 ☐
c. Effective use of language (e.g., sentence variety, word choice, tone)	1 ☐	2 ☐	3 ☐	4 ☐
d. Mechanics (e.g., spelling, capitalization, punctuation, grammar) and conventions, paragraphing, using appropriate formats, such as subheadings and quotations)	1 ☐	2 ☐	3 ☐	4 ☐
e. For text-based writing, the accuracy or appropriateness of their references to texts	1 ☐	2 ☐	3 ☐	4 ☐
f. Quality of thinking	1 ☐	2 ☐	3 ☐	4 ☐
g. Demonstration of content knowledge or understanding of text	1 ☐	2 ☐	3 ☐	4 ☐

D3. How often do you ask your students to do the following when you ask them to write about something?

SELECT ONE RESPONSE PER ROW

	NEVER OR HARDLY EVER	SOMETIMES	OFTEN	VERY OFTEN OR ALWAYS
a. Plan their writing	1 ☐	2 ☐	3 ☐	4 ☐
b. Analyze strong writing models and mentor texts	1 ☐	2 ☐	3 ☐	4 ☐
c. Create criteria to evaluate their writing	1 ☐	2 ☐	3 ☐	4 ☐
d. Provide critiques to peers	1 ☐	2 ☐	3 ☐	4 ☐
e. Define their purpose and audience	1 ☐	2 ☐	3 ☐	4 ☐
f. Make a formal outline before they write	1 ☐	2 ☐	3 ☐	4 ☐
g. Write more than one draft of a paper	1 ☐	2 ☐	3 ☐	4 ☐
h. Use a dictionary, thesaurus or encyclopedia	1 ☐	2 ☐	3 ☐	4 ☐
i. Check for proper spelling, grammar, and punctuation themselves	1 ☐	2 ☐	3 ☐	4 ☐
j. Use of a peer review protocol	1 ☐	2 ☐	3 ☐	4 ☐
k. Assess their own writing by using a specific rubric	1 ☐	2 ☐	3 ☐	4 ☐
l. Assess the writing of other students by using rubrics	1 ☐	2 ☐	3 ☐	4 ☐
m. Present or share their writing to the whole class or others outside the class	1 ☐	2 ☐	3 ☐	4 ☐
n. Write for social networking, blogs or wikis	1 ☐	2 ☐	3 ☐	4 ☐
o. Write citing evidence or information from a text they have read	1 ☐	2 ☐	3 ☐	4 ☐
p. Create videos or webcasts	1 ☐	2 ☐	3 ☐	4 ☐

D4. <u>In the past two weeks of regular instruction</u>, how often did you ask your students to write for each of the following purposes?

SELECT ONE RESPONSE PER ROW

	NEVER	ONCE OR TWICE A WEEK	ALMOST EVERY DAY	EVERYDAY	MORE THAN ONCE A DAY
a. To inform (e.g., provide information about a topic or steps in a process)	1 ☐	2 ☐	3 ☐	4 ☐	4 ☐
b. To argue or present an opinion (e.g., to support a claim with evidence from sources)	1 ☐	2 ☐	3 ☐	4 ☐	4 ☐
c. To convey experience in narrative form (write about personal or imagined experiences)	1 ☐	2 ☐	3 ☐	4 ☐	4 ☐

D5. **<u>In the past two weeks of regular instruction</u>, how often did you ask your students to write the following?**

SELECT ONE RESPONSE PER ROW

	NEVER	ONCE A WEEK	2 TO 4 TIMES A WEEK	ALMOST EVERY DAY	EVERY DAY
a. Short responses to questions or note-taking, where students copy or record information	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
b. Short responses to questions, such as "think tasks" or other tasks that promote students' thinking	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
c. Thinking task, such as graphic organizers, that capture students' thoughts on the text and relevant evidence	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
d. Writing that cites evidence or information from a text they have read	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
e. Journal entries	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
f. Letters	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
g. Speeches	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
h. Narratives (non-fiction)	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
i. Multi-step instructions or explanations of how to perform a task	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
j. Fictional stories	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
k. Poems	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
l. Book reviews	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
m. Literature reviews	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
n. Literary analysis essays	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
o. Personal essays	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
p. Informational essays	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
q. Position papers or persuasive essays	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
r. Argumentative essays	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
s. On-demand responses to writing prompts	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

D6. <u>In the past two weeks</u>, how often did you ask your students to write for each of the following audiences?

SELECT ONE RESPONSE PER ROW

	NEVER	ONCE A WEEK	2 TO 4 TIMES A WEEK	ALMOST EVERY DAY
a. Family members	1 ☐	2 ☐	3 ☐	4 ☐
b. School officials	1 ☐	2 ☐	3 ☐	4 ☐
c. Teacher	1 ☐	2 ☐	3 ☐	4 ☐
d. Other Students	1 ☐	2 ☐	3 ☐	4 ☐
e. Community members	1 ☐	2 ☐	3 ☐	4 ☐
f. Community organizations, government officials, or businesses	1 ☐	2 ☐	3 ☐	4 ☐

D6g. In the past two weeks, were there any other audiences you asked your students to write for?

- ❍ Yes ..... 1 GO TO D6gspec
- ❍ No ..... 0 GO TO EL TPP CHECK
- ❍ Don't know ..... 0 GO TO EL TPP CHECK

NO RESPONSE ..... M GO TO EL TPP CHECK

D6gspec For what other audiences did you ask your students to write?

ALLOW 280

D6grate In the past two weeks, how often did you ask your students to write for these other audiences?

Select one only

- ❍ Never 1
- ❍ Once a week 2
- ❍ Two or four times a week 3
- ❍ Almost every day 4

EL TPP CHECK: IF YOUR SCHOOL <u>IS NOT</u> PARTICIPATING IN EL TPP, GO TO SECTION F ON PAGE 29.

E. EL EDUCATION PROFESSIONAL DEVELOPMENT

This section collects information about your experiences with EL Education coaches, EL Education institutes and EL Education web supports, including Professional Learning Packs and webinars.

E1. Since [REFERENCE DATE], have you met with an EL Education coach in a small group or individually, in person or through online video coaching?

- ❍ Yes 1
- ❍ No ..... 0 GO TO E5
- ❍ Don't know ..... 0 GO TO E
  NO RESPONSE ..... M GO TO E5

E2. Since [REFERENCE DATE], how many times have you met with an EL Education coach either in-person or through online video coaching? Do not include any scheduled meetings that have not yet occurred.

Select one only

- ❍ 1 to 2 times 1
- ❍ 3-5 times 2
- ❍ 6 to 10 times 3
- ❍ 11 or more times 4

E2a. Since [REFERENCE DATE], how many times have you met <u>in-person</u> with an EL Education coach? Do not include any scheduled meetings that have not yet occurred.

Select one only

- ❍ 1 to 2 times 1
- ❍ 3-5 times 2
- ❍ 6 to 10 times 3
- ❍ 11 or more times 4

E2b. Since [REFERENCE DATE], how many times have you had an <u>online</u> video meeting with an EL Education coach? Do not include any scheduled meetings that have not yet occurred.

Select one only

- ❍ 1 to 2 times 1
- ❍ 3-5 times 2
- ❍ 6 to 10 times 3
- ❍ 11 or more times 4

E3. Thinking about your meetings with your coach since [REFERENCE DATE], about how long is a typical meeting with your EL Education coach?

Select one only

- ❍ Less than 1 hour 1
- ❍ 1 to less than 3 hours 2
- ❍ 3 or more hours 3

E3a. In total, about how many hours have you worked with your EL Education coach since [REFERENCE DATE]? Your best estimate is fine.

|___|___|___| NUMBER OF HOURS MET WITH EL EDUCATION COACH SINCE [REFERENCE DATE]

E3b. In total, about how many hours did you meet <u>in-person</u> with your EL Education coach since [REFERENCE DATE]? Your best estimate is fine.

|___|___|___| NUMBER OF HOURS MET <u>IN-PERSON</u> WITH EL EDUCATION COACH SINCE [REFERENCE DATE]

E3c. In total, about how many hours did you participate in <u>online</u> video meetings with your EL Education coach since [REFERENCE DATE]? Your best estimate is fine.

|___|___|___| NUMBER OF HOURS MET <u>ONLINE</u> MEETINGS WITH EL EDUCATION COACH SINCE [REFERENCE DATE]

E4a1. Since [REFERENCE DATE], did you participate in an in-person coaching meeting that covered the topic of "implementing a Common Core-aligned curriculum?"

- ❍ Yes ..... 1 GO TO E4b1
- ❍ No ..... 0 GO TO E4a2

NO RESPONSE ..... M GO TO E4a2

E4b1. How useful to you was the in-person coaching on "<u>implementing a Common Core-aligned curriculum?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a2. Since [REFERENCE DATE], did you participate in an in-person coaching meeting that covered the topic of "engaging students in discussion and thinking protocols?"

- ❍ Yes ..... 1 GO TO E4b2
- ❍ No ..... 0 GO TO E4a3

NO RESPONSE ..... M GO TO E4a3

E4b2. How useful to you was the in-person coaching on "<u>engaging students in discussion and thinking protocols?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a3. Since [REFERENCE DATE], did you participate in an in-person coaching meeting that covered the topic of "helping students read complex text through close reading and other strategies?"

- ❍ Yes ..... 1 GO TO E4b3
- ❍ No ..... 0 GO TO E4a4

NO RESPONSE ..... M GO TO E4a4

E4b3. How useful to you was the in-person coaching on "<u>helping students read complex text through close reading and other strategies?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a4. Since [REFERENCE DATE], did you participate in an in-person coaching meeting that covered the topic of "helping students select and provide evidence for their thinking?"

- ❍ Yes ..... 1 GO TO E4b4
- ❍ No ..... 0 GO TO E4a5

NO RESPONSE ..... M GO TO E4a5

E4b4. How useful to you was the in-person coaching on "<u>helping students select and provide evidence for their thinking?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a5. Since [REFERENCE DATE], did you participate in an in-person coaching meeting that covered the topic of "helping students learn vocabulary?"

- ❍ Yes ..... 1 GO TO E4b5
- ❍ No ..... 0 GO TO E4a6

NO RESPONSE ..... M GO TO E4a6

E4b5. How useful to you was the in-person coaching on "<u>helping students learn vocabulary?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a6. Since [REFERENCE DATE], did you participate in an in-person coaching meeting that covered the topic of "engaging students in the analysis of models of writing?"

- ❍ Yes ..... 1 GO TO E4b6
- ❍ No ..... 0 GO TO E4a7

NO RESPONSE ..... M GO TO E4a7

E4b6. How useful to you was the in-person coaching on "<u>engaging students in the analysis of models of writing?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a7. Since [REFERENCE DATE], did you participate in an in-person coaching meeting that covered the topic of "using curriculum-based and other data to drive instruction (student's writing and assessments)?"

- ❍ Yes ..... 1 GO TO E4b7
- ❍ No ..... 0 GO TO E4a8

NO RESPONSE ..... M GO TO E4a8

E4b7. How useful to you was the in-person coaching on "<u>using curriculum-based and other data to drive instruction (student's writing and assessments)?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a8. Since [REFERENCE DATE], did you participate in an in-person coaching meeting that covered the topic of "developing writing from sources?"

❍ Yes	1	GO TO E4b8
❍ No	0	GO TO E4a9
NO RESPONSE	M	GO TO E4a9

E4b8. How useful to you was the in-person coaching on "<u>developing writing from sources?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a9. Since [REFERENCE DATE], did you participate in an in-person coaching meeting that covered the topic of "supporting English Language Learners?"

❍ Yes	1	GO TO E4b9
❍ No	0	GO TO E4a10
NO RESPONSE	M	GO TO E4a10

E4b9. How useful to you was the in-person coaching on "<u>supporting English Language Learners?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a10. Since [REFERENCE DATE], did you participate in an in-person coaching meeting that covered the topic of "supporting Special Education Students?"

❍ Yes	1	GO TO E4b10
❍ No	0	GO TO E4a11
NO RESPONSE	M	GO TO E4a11

E4b10. How useful to you was the in-person coaching on "<u>supporting Special Education Students?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a11. Since [REFERENCE DATE], did you participate in an in-person coaching meeting that covered the topic of "using learning targets?"

❍ Yes	1	GO TO E4b11
❍ No	0	GO TO E4a12
NO RESPONSE	M	GO TO E4a12

E4b11. How useful to you was the in-person coaching on "<u>using learning targets?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a12. Since [REFERENCE DATE], did you participate in an in-person coaching meeting that covered the topic of "using formative assessment strategies?"

❍ Yes	1	GO TO E4b12
❍ No	0	GO TO E4a13
NO RESPONSE	M	GO TO E4a13

E4b12. How useful to you was the in-person coaching on "<u>using formative assessment strategies?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a13. Since [REFERENCE DATE], did you participate in an in-person coaching meeting that covered the topic of "supporting students in tracking their learning and setting goals?"

❍ Yes	1	GO TO E4b13
❍ No	0	GO TO E4a14
NO RESPONSE	M	GO TO E4a14

E4b13. How useful to you was the in-person coaching on "<u>supporting students in tracking their learning and setting goals?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a14. Since [REFERENCE DATE], did you participate in an in-person coaching meeting that covered the topic of "providing peer critique and other feedback?"

❍ Yes	1	GO TO E4b14
❍ No	0	GO TO E4a15
NO RESPONSE	M	GO TO E4a15

E4b14. How useful to you was the in-person coaching on "providing peer critique and other feedback?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a15. Since [REFERENCE DATE], did you participate in an in-person coaching meeting that covered the topic of "teachers and students co-constructing criteria for success (academic and behavioral)?"

❍ Yes	1	GO TO E4b15
❍ No	0	GO TO E4a16
NO RESPONSE	M	GO TO E4a16

E4b15. How useful to you was the in-person coaching on "teachers and students co-constructing criteria for success (academic and behavioral)?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a16. Since [REFERENCE DATE], did you participate in an in-person coaching meeting that covered the topic of "managing an active classroom where students are active participants/leaders?"

❍ Yes	1	GO TO E4b16
❍ No	0	GO TO E4a17
NO RESPONSE	M	GO TO E4a17

E4b16. How useful to you was the in-person coaching on "managing an active classroom where students are active participants/leaders?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

The next questions are about online or video-based coaching.

E4a17. Since [REFERENCE DATE], did you participate in an online coaching meeting that covered the topic of "implementing a Common Core-aligned curriculum?"

❍ Yes	1	GO TO E4b17
❍ No	0	GO TO E4a18
NO RESPONSE	M	GO TO E4a18

E4b17. How useful to you was the online coaching on "<u>implementing a Common Core-aligned curriculum?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a18. Since [REFERENCE DATE], did you participate in an online coaching meeting that covered the topic of "engaging students in discussion and thinking protocols?"

❍ Yes	1	GO TO E4b18
❍ No	0	GO TO E4a19
NO RESPONSE	M	GO TO E4a19

E4b18. How useful to you was the online coaching on "<u>engaging students in discussion and thinking protocols?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a19. Since [REFERENCE DATE], did you participate in an online coaching meeting that covered the topic of "helping students read complex text through close reading and other strategies?"

❍ Yes	1	GO TO E4b19
❍ No	0	GO TO E4a20
NO RESPONSE	M	GO TO E4a20

E4b19. How useful to you was the online coaching on "<u>helping students read complex text through close reading and other strategies?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a20. Since [REFERENCE DATE], did you participate in an online coaching meeting that covered the topic of "helping students select and provide evidence for their thinking?"

❍ Yes	1	GO TO E4b20
❍ No	0	GO TO E4a21
NO RESPONSE	M	GO TO E4a21

E4b20. How useful to you was the online coaching on "<u>helping students select and provide evidence for their thinking?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a21. Since [REFERENCE DATE], did you participate in an online coaching meeting that covered the topic of "helping students learn vocabulary?"

❍ Yes	1	GO TO E4b21
❍ No	0	GO TO E4a22
NO RESPONSE	M	GO TO E4a22

E4b21. How useful to you was the online coaching on "<u>helping students learn vocabulary?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a22. Since [REFERENCE DATE], did you participate in an online coaching meeting that covered the topic of "engaging students in the analysis of models of writing?"

❍ Yes	1	GO TO E4b22
❍ No	0	GO TO E4a23
NO RESPONSE	M	GO TO E4a23

E4b22. How useful to you was the online coaching on "<u>engaging students in the analysis of models of writing?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a23. Since [REFERENCE DATE], did you participate in an online coaching meeting that covered the topic of "using curriculum-based and other data to drive instruction (student's writing and assessments)?"

❍ Yes	1	GO TO E4b23
❍ No	0	GO TO E4a24
NO RESPONSE	M	GO TO E4a24

E4b23. How useful to you was the online coaching on "<u>using curriculum-based and other data to drive instruction (student's writing and assessments)?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a24. Since [REFERENCE DATE], did you participate in an online coaching meeting that covered the topic of "developing writing from sources?"

❍ Yes	1	GO TO E4b24
❍ No	0	GO TO E4a25
NO RESPONSE	M	GO TO E4a25

E4b24. How useful to you was the online coaching on "<u>developing writing from sources?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a25. Since [REFERENCE DATE], did you participate in an online coaching meeting that covered the topic of "supporting English Language Learners?"

❍ Yes	1	GO TO E4b25
❍ No	0	GO TO E4a26
NO RESPONSE	M	GO TO E4a26

E4b25. How useful to you was the online coaching on "supporting English Language Learners?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a26. Since [REFERENCE DATE], did you participate in an online coaching meeting that covered the topic of "supporting Special Education Students?"

❍ Yes	1	GO TO E4b26
❍ No	0	GO TO E4a27
NO RESPONSE	M	GO TO E4a127

E4b26. How useful to you was the online coaching on "supporting Special Education Students?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a27. Since [REFERENCE DATE], did you participate in an online coaching meeting that covered the topic of "using learning targets?"

❍ Yes	1	GO TO E4b27
❍ No	0	GO TO E4a28
NO RESPONSE	M	GO TO E4a28

E4b27. How useful to you was the online coaching on "using learning targets?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a28. Since [REFERENCE DATE], did you participate in an online coaching meeting that covered the topic of "using formative assessment strategies?"

❍ Yes	1	GO TO E4b28
❍ No	0	GO TO E4a29
NO RESPONSE	M	GO TO E4a29

E4b28. How useful to you was the online coaching on "<u>using formative assessment strategies?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a29. Since [REFERENCE DATE], did you participate in an online coaching meeting that covered the topic of "supporting students in tracking their learning and setting goals?"

- ❍ Yes ..... 1 GO TO E4b29
- ❍ No ..... 0 GO TO E4a30

NO RESPONSE ..... M GO TO E4a30

E4b29. How useful to you was the online coaching on "<u>supporting students in tracking their learning and setting goals?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a30. Since [REFERENCE DATE], did you participate in an online coaching meeting that covered the topic of "providing peer critique and other feedback?"

- ❍ Yes ..... 1 GO TO E4b30
- ❍ No ..... 0 GO TO E4a31

NO RESPONSE ..... M GO TO E4a31

E4b30. How useful to you was the online coaching on "<u>providing peer critique and other feedback?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E4a31. Since [REFERENCE DATE], did you participate in an online coaching meeting that covered the topic of "teachers and students co-constructing criteria for success (academic and behavioral)?"

- ❍ Yes ..... 1 GO TO E4b31
- ❍ No ..... 0 GO TO E4a32

NO RESPONSE ..... M GO TO E4a32

E4b31. How useful to you was the online coaching on "<u>teachers and students co-constructing criteria for success (academic and behavioral)?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3

❍ Very useful 4

E4a32. Since [REFERENCE DATE], did you participate in an online coaching meeting that covered the topic of "managing an active classroom where students are active participants/leaders?"

❍ Yes	1	GO TO E4b32
❍ No	0	GO TO E5
NO RESPONSE	M	GO TO E5

E4b32. How useful to you was the online coaching on "<u>managing an active classroom where students are active participants/leaders?</u>"

Select one only

❍ Not at all useful 1
❍ Somewhat useful 2
❍ Useful 3
❍ Very useful 4

E5. Since [REFERENCE DATE], have you attended any EL Education <u>institutes</u>?

Institutes are formal, PD sessions, for one to two days, with an EL Education coach.

❍ Yes	1	
❍ No	0	GO TO E8
NO RESPONSE	M	GO TO E8

E6. Since [REFERENCE DATE], how many EL Education <u>institutes</u> have you attended so far? Do not include any scheduled institutes that have not yet occurred.

Institutes are formal, PD sessions, for one to two days, with an EL Education coach.

Select one only

❍ 1 1
❍ 2 2
❍ 3 3
❍ 4 4
❍ 5 or more 5

E6a. Since [REFERENCE DATE] how many hours in total have you attended EL Education institutes this school year? Your best estimate is fine.

Institutes are formal, PD sessions, for one to two days, with an EL Education coach.

|___|___| TOTAL NUMBER OF HOURS ATTENDED EL EDUCATION INSTITUTES SINCE [REFERENCE DATE]

E7a1. Since [REFERENCE DATE], did you attend an EL Education institute that covered the topic of "implementing a Common Core-aligned curriculum?"

❍ Yes	1	GO TO E7b1
❍ No	0	GO TO E7a2
NO RESPONSE	M	GO TO E7a2

E7b1. How useful to you was the institute on "<u>implementing a Common Core-aligned curriculum?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E7a2. Since [REFERENCE DATE], did you attend an EL Education institute that covered the topic of "engaging students in discussion and thinking protocols?"

❍ Yes	1	GO TO E7b2
❍ No	0	GO TO E7a3
NO RESPONSE	M	GO TO E7a3

E7b2. How useful to you was the institute on "<u>engaging students in discussion and thinking protocols?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E7a3. Since [REFERENCE DATE], did you attend an EL Education institute that covered the topic of "helping students read complex text through close reading and other strategies?"

❍ Yes	1	GO TO E7b3
❍ No	0	GO TO E7a4
NO RESPONSE	M	GO TO E7a4

E7b3. How useful to you was the institute on "<u>helping students read complex text through close reading and other strategies?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E7a4. Since [REFERENCE DATE], did you attend an EL Education institute that covered the topic of "helping students select and provide evidence for their thinking?"

❍ Yes	1	GO TO E7b4
❍ No	0	GO TO E7a5
NO RESPONSE	M	GO TO E7a5

E7b4. How useful to you was the institute on "helping students select and provide evidence for their thinking?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E7a5. Since [REFERENCE DATE], did you attend an EL Education institute that covered the topic of "helping students learn vocabulary?"

❍ Yes	1	GO TO E7b5
❍ No	0	GO TO E7a6
NO RESPONSE	M	GO TO E7a6

E7b5. How useful to you was the institute on "helping students learn vocabulary?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E7a6. Since [REFERENCE DATE], did you attend an EL Education institute that covered the topic of "engaging students in the analysis of models of writing?"

❍ Yes	1	GO TO E7b6
❍ No	0	GO TO E7a7
NO RESPONSE	M	GO TO E7a7

E7b6. How useful to you was the institute on "engaging students in the analysis of models of writing?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E7a7. Since [REFERENCE DATE], did you attend an EL Education institute that covered the topic of "using curriculum-based and other data to drive instruction (student's writing and assessments)?"

- ❍ Yes ..... 1 GO TO E7b7
- ❍ No ..... 0 GO TO E7a8

NO RESPONSE ..... M GO TO E7a8

E7b7. How useful to you was the institute on "<u>using curriculum-based and other data to drive instruction (student's writing and assessments)?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E7a8. Since [REFERENCE DATE], did you attend an EL Education institute that covered the topic of "developing writing from sources?"

- ❍ Yes ..... 1 GO TO E7b8
- ❍ No ..... 0 GO TO E7a9

NO RESPONSE ..... M GO TO E7a9

E7b8. How useful to you was the institute on "<u>developing writing from sources?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E7a9. Since [REFERENCE DATE], did you attend an EL Education institute that covered the topic of "supporting English Language Learners?"

- ❍ Yes ..... 1 GO TO E7b9
- ❍ No ..... 0 GO TO E7a10

NO RESPONSE ..... M GO TO E7a10

E7b9. How useful to you was the institute on "<u>supporting English Language Learners?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E7a10. Since [REFERENCE DATE], did you attend an EL Education institute that covered the topic of "supporting Special Education Students?"

- ❍ Yes ..... 1 GO TO E7b10
- ❍ No ..... 0 GO TO E7a11

NO RESPONSE ..... M GO TO E7a11

E7b10. How useful to you was the institute on "<u>supporting Special Education Students?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E7a11. Since [REFERENCE DATE], did you attend an EL Education institute that covered the topic of "using learning targets?"

- ❍ Yes ..... 1 GO TO E7b11
- ❍ No ..... 0 GO TO E7a12

NO RESPONSE ..... M GO TO E7a12

E7b11. How useful to you was the institute on "<u>using learning targets?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E7a12. Since [REFERENCE DATE], did you attend an EL Education institute that covered the topic of "using formative assessment strategies?"

- ❍ Yes ..... 1 GO TO E7b12
- ❍ No ..... 0 GO TO E7a13

NO RESPONSE ..... M GO TO E7a13

E7b12. How useful to you was the institute on "<u>using formative assessment strategies?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E7a13. Since [REFERENCE DATE], did you attend an EL Education institute that covered the topic of "supporting students in tracking their learning and setting goals?"

- ❍ Yes ..... 1 GO TO E7b13
- ❍ No ..... 0 GO TO E7a14

NO RESPONSE ..... M GO TO E7a14

E7b13. How useful to you was the institute on "<u>supporting students in tracking their learning and setting goals?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E7a14. Since [REFERENCE DATE], did you attend an EL Education institute that covered the topic of "providing peer critique and other feedback?"

- ❍ Yes ..... 1 GO TO E7b14
- ❍ No ..... 0 GO TO E7a15

NO RESPONSE ..... M GO TO E7a15

E7b14. How useful to you was the institute on "<u>providing peer critique and other feedback?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E7a15. Since [REFERENCE DATE], did you attend an EL Education institute that covered the topic of "teachers and students co-constructing criteria for success (academic and behavioral)?"

- ❍ Yes ..... 1 GO TO E7b15
- ❍ No ..... 0 GO TO E7a16

NO RESPONSE ..... M GO TO E7a16

E7b15. How useful to you was the institute on "<u>teachers and students co-constructing criteria for success (academic and behavioral)?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E7a16. Since [REFERENCE DATE], did you attend an EL Education institute that covered the topic of "managing an active classroom where students are active participants/leaders?"

❍ Yes	1	GO TO E7b16
❍ No	0	GO TO E8
NO RESPONSE	M	GO TO E8A

E7b16. How useful to you was the institute on "<u>managing an active classroom where students are active participants/leaders?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E8. Have you accessed any online EL Education supports?

Online supports include: webinars, Professional Learning Packs (PLPs), online facilitated conversations, online professional development materials, and online community of practice. DO NOT include online coaching.

PLPs are collections of materials designed for self-study, small group professional development, or large group professional development.

- ❍ Yes 1
- ❍ No 0

IF you answered "NO" to E1, E5, E8, then go to E13. Otherwise, continue to E12

E9. Since [REFERENCE DATE], **about how many times have you participated in EL Education webinars? Do not include any scheduled webinars that have not occurred.**

Webinars are formal online, hour-long presentations, facilitated by EL Education coaches.

Select one only

- ❍ 1 to 2 1
- ❍ 3 to 5 2
- ❍ 6 to 10 3
- ❍ 11 or more times 4

E10. Since [REFERENCE DATE], **about how many times have you used Professional Learning Packs (PLPs)?**

PLPs are collections of materials designed for self-study, small group professional development, or large group professional development.

Select one only

- ❍ 1 to 2 1
- ❍ 3 to 5 2
- ❍ 6 to 10 3
- ❍ 11 or more times 4

E10a. Since [REFERENCE DATE], how many hours in total have you spent accessing EL Education online supports? Your best estimate is fine.

Online supports include: webinars, Professional Learning Packs (PLPS), online facilitated conversations, online professional development materials, and online community of practice.

DO NOT include online coaching.

|___|___|___| NUMBER OF HOURS ACCESSED ONLINE SUPPORT

E11a1. Since [REFERENCE DATE], did you attend an EL Education webinar or access EL Education online materials that covered the topic of "implementing a Common Core-aligned curriculum?"

❍ Yes	1	GO TO E11b1
❍ No	0	GO TO E11a2
NO RESPONSE	M	GO TO E11a2

E11b1. How useful to you was the online support on "<u>implementing a Common Core-aligned curriculum?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E11a2. Since [REFERENCE DATE], did you attend an EL Education webinar or access EL Education online materials that covered the topic of "engaging students in discussion and thinking protocols?"

❍ Yes	1	GO TO E11b2
❍ No	0	GO TO E11a3
NO RESPONSE	M	GO TO E11a3

E11b2. How useful to you was the online support on "<u>engaging students in discussion and thinking protocols?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E11a3. Since [REFERENCE DATE], did you attend an EL Education webinar or access EL Education online materials that covered the topic of "helping students read complex text through close reading and other strategies?"

❍ Yes	1	GO TO E11b3
❍ No	0	GO TO E11a4
NO RESPONSE	M	GO TO E11a4

E11b3. How useful to you was the online support on "<u>helping students read complex text through close reading and other strategies?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E11a4. Since [REFERENCE DATE], did you attend an EL Education webinar or access EL Education online materials that covered the topic of "helping students select and provide evidence for their thinking?"

❍ Yes	1	GO TO E11b4
❍ No	0	GO TO E11a5
NO RESPONSE	M	GO TO E11a5

E11b4. How useful to you was the online support on "helping students select and provide evidence for their thinking?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E11a5. Since [REFERENCE DATE], did you attend an EL Education webinar or access EL Education online materials that covered the topic of "helping students learn vocabulary?"

❍ Yes	1	GO TO E11b5
❍ No	0	GO TO E11a6
NO RESPONSE	M	GO TO E11a6

E11b5. How useful to you was the online support on "helping students learn vocabulary?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E11a6. Since [REFERENCE DATE], did you attend an EL Education webinar or access EL Education online materials that covered the topic of "engaging students in the analysis of models of writing?"

❍ Yes	1	GO TO E11b6
❍ No	0	GO TO E11a7
NO RESPONSE	M	GO TO E11a7

E11b6. How useful to you was the online support on "engaging students in the analysis of models of writing?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E11a7. Since [REFERENCE DATE], did you attend an EL Education webinar or access EL Education online materials that covered the topic of "using curriculum-based and other data to drive instruction (student's writing and assessments)?"

❍ Yes	1	GO TO E11b7
❍ No	0	GO TO E11a8
NO RESPONSE	M	GO TO E11a8

E11b7. How useful to you was the online support on "<u>using curriculum-based and other data to drive instruction (student's writing and assessments)?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E11a8. Since [REFERENCE DATE], did you attend an EL Education webinar or access EL Education online materials that covered the topic of "developing writing from sources?"

❍ Yes	1	GO TO E11b8
❍ No	0	GO TO E11a9
NO RESPONSE	M	GO TO E11a9

E11b8. How useful to you was the online support on "<u>developing writing from sources?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E11a9. Since [REFERENCE DATE], did you attend an EL Education webinar or access EL Education online materials that covered the topic of "supporting English Language Learners?"

❍ Yes	1	GO TO E11b9
❍ No	0	GO TO E11a10
NO RESPONSE	M	GO TO E11a10

E11b9. How useful to you was the online support on "supporting English Language Learners?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E11a10. Since [REFERENCE DATE], did you attend an EL Education webinar or access EL Education online materials that covered the topic of "supporting Special Education Students?"

❍ Yes	1	GO TO E11b10
❍ No	0	GO TO E11a11
NO RESPONSE	M	GO TO E11a11

E11b10. How useful to you was the online support on "supporting Special Education Students?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E11a11. Since [REFERENCE DATE], did you attend an EL Education webinar or access EL Education online materials that covered the topic of "using learning targets?"

❍ Yes	1	GO TO E11b11
❍ No	0	GO TO E11a12
NO RESPONSE	M	GO TO E11a12

E11b11. How useful to you was the online support on "using learning targets?"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E11a12. Since [REFERENCE DATE], did you attend an EL Education webinar or access EL Education online materials that covered the topic of "using formative assessment strategies?"

❍ Yes	1	GO TO E11b12
❍ No	0	GO TO E11a13
NO RESPONSE	M	GO TO E11a13

E11b12. How useful to you was the online support on "<u>using formative assessment strategies?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E11a13. Since [REFERENCE DATE], did you attend an EL Education webinar or access EL Education online materials that covered the topic of "supporting students in tracking their learning and setting goals?"

❍ Yes	1	GO TO E11b13
❍ No	0	GO TO E11a14
NO RESPONSE	M	GO TO E11a14

E11b13. How useful to you was the online support on "<u>supporting students in tracking their learning and setting goals?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E11a14. Since [REFERENCE DATE], did you attend an EL Education webinar or access EL Education online materials that covered the topic of "providing peer critique and other feedback?"

❍ Yes	1	GO TO E11b14
❍ No	0	GO TO E11a15
NO RESPONSE	M	GO TO E11a15

E11b14. How useful to you was the online support on "<u>providing peer critique and other feedback?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E11a15. Did you attend an EL Education webinar or access EL Education online materials that covered the topic of "teachers and students co-constructing criteria for success (academic and behavioral)?"

- ❍ Yes ..... 1 GO TO E11b15
- ❍ No ..... 0 GO TO E11a16

NO RESPONSE ..... M GO TO E11a16

E11b15. Since [REFERENCE DATE], did useful to you was the online support on "<u>teachers and students co-constructing criteria for success (academic and behavioral)?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

E11a16. Since [REFERENCE DATE], did you attend an EL Education webinar or access EL Education online materials that covered the topic of "managing an active classroom where students are active participants/leaders?"

- ❍ Yes ..... 1 → GO TO E11b16
- ❍ No ..... 0 →

NO RESPONSE ..... M → GO TO PROGRAMMER INSTRUCTION

E11b16. How useful to you was the online support on "<u>managing an active classroom where students are active participants/leaders?</u>"

Select one only

- ❍ Not at all useful 1
- ❍ Somewhat useful 2
- ❍ Useful 3
- ❍ Very useful 4

IF YOU HAVE NOT PARTICIPATED IN ANY PROFESSIONAL DEVELOPMENT (PD),

YOU MARKED "NO" TO QUESTIONS E1, E5, AND E8,

THEN GO TO SECTION F ON PAGE 29

E12. Taking into account all PD you participated in, to what degree did each of the professional development topics below impact your classroom practice?

SELECT ONE RESPONSE PER ROW

	NOT AT ALL (I do not use the techniques/ strategies.)	**SOMEWHAT** (I use some of the techniques/ strategies.)	**A GOOD DEAL** (I use many of the techniques/ strategies.)	**EXTENSIVELY** (My teaching changed fundamentally.)	**NA** (Did not receive PD on this topic)
a. Implementing EL Education's Common Core-aligned curriculum	0 ☐	1 ☐	2 ☐	3 ☐	n ☐
b. Engaging students in discussion and thinking protocols	0 ☐	1 ☐	2 ☐	3 ☐	n ☐
c. Helping students read complex text through close reading and other strategies	0 ☐	1 ☐	2 ☐	3 ☐	n ☐
d. Helping students select and provide evidence for their thinking	0 ☐	1 ☐	2 ☐	3 ☐	n ☐
e. Helping students learn vocabulary	0 ☐	1 ☐	2 ☐	3 ☐	n ☐
f. Engaging students in the analysis of models of writing	0 ☐	1 ☐	2 ☐	3 ☐	n ☐
g. Using curriculum-based and other data to drive instruction (student's writing and assessments)	0 ☐	1 ☐	2 ☐	3 ☐	n ☐
h. Developing writing from sources	0 ☐	1 ☐	2 ☐	3 ☐	n ☐
i. Supporting English Language Learners	0 ☐	1 ☐	2 ☐	3 ☐	n ☐
j. Supporting Special Education students	0 ☐	1 ☐	2 ☐	3 ☐	n ☐
k. Using learning targets	0 ☐	1 ☐	2 ☐	3 ☐	n ☐
l. Using formative assessment strategies	0 ☐	1 ☐	2 ☐	3 ☐	n ☐
m. Supporting students in tracking their learning and setting goals	0 ☐	1 ☐	2 ☐	3 ☐	n ☐
n. Providing peer critique and other feedback	0 ☐	1 ☐	2 ☐	3 ☐	n ☐
o. Teachers and students co-constructing criteria for success (academic and/or behavioral)	0 ☐	1 ☐	2 ☐	3 ☐	n ☐
p. Managing an active classroom where students are active participants/leaders	0 ☐	1 ☐	2 ☐	3 ☐	n ☐

E12q Were there other professional development topics that had an impact on your classroom practice?

- ❍ Yes ..... 1 GO TO E12qspec
- ❍ No ..... 0 GO TO E13
- ❍ Don't know ..... 0 GO TO E13

NO RESPONSE ..... M GO TO E13

E12qspec What other professional development topics had an impact on your classroom practice?

E12qrate To what degree have these other professional development topics made an impact on your classroom practice?

Select one only

- ❍ Not at all (I do not use the techniques/strategies.) 1
- ❍ Somewhat (I use dome of the techniques/strategies) 2
- ❍ A good deal (I sue many o fthe techniques/strategies.) 3
- ❍ Extensively (My teaching changed fundamentally) 4

E13. To what extent have you implemented the EL Education curriculum?

Select one only

- ❍ Total implementation 1
- ❍ Nearly total implementation (implemented 75 percent of the curriculum) 2
- ❍ Partial implementation (implemented 50 percent of the curriculum) 3
- ❍ Very little implementation (implemented 25 percent of the curriculum) 4
- ❍ No implementation ..... 4 GO TO E15

E14. Which modules of the EL Education curriculum did you implement?

SELECT ALL THAT APPLY

Third Grade ELA

1 ☐ Module 1: The Power of Reading

2 ☐ Module 2A: Researching to Build Knowledge and Teaching Others: Adaptations and the Wide World of Frogs

3 ☐ Module 2B: Researching to Build Knowledge and Teaching Others: Connecting Literacy and Informational Texts to Study Culture "Then and Now"

4 ☐ Module 3A: Analyzing Narrative and Supporting Opinions: A Study of Peter Pan

5 ☐ Module 3B: Analyzing and Narrative Supporting Opinions: Wolves in Fiction and Fact

6 ☐ Module 4: Gathering Evidence and Speaking to Others: The Role of Freshwater Around the World

Fourth Grade ELA

7 ☐ Module 1A: Becoming a Close Reader and Writing to Learn: Oral Tradition, Symbolism, Building Community

8 ☐ Module 1B: Reading Closely and Writing to Learn: Poetry, Poets, and Becoming Writers

9 ☐ Module 2A: Researching to Build Knowledge and Teaching Others: Interdependent Roles in Colonial Times

10 ☐ Module 2B: Researching to Build Knowledge and Teaching Others: Animal Defense Mechanisms

11 ☐ Module 3A: Considering Perspectives and Supporting Opinions: Simple Machines: Force and Motion

12 ☐ Module 3B: Considering Perspectives and Supporting Opinions: Perspectives on the American Revolution

13 ☐ Module 4: Considering Perspectives and Supporting Opinions: Suffrage and the Leadership of Susan B. Anthony

Fifth Grade ELA

14 ☐ Module 1: Becoming a Close Reader and Writing to Learn: Stories of Human Rights

15 ☐ Module 2A: Researching to Build Knowledge and Teaching Others: Biodiversity in Rainforests of the Western Hemisphere

16 ☐ Module 2B: Researching to Build Knowledge and Teaching Others: Inventions that Changed People's Lives

17 ☐ Module 3A: Considering Perspectives and Supporting Opinions: Sports and Athletes' Impact on Culture

18 ☐ Module 3B: Considering Perspectives and Supporting Opinions: Balancing Competing Needs in Canada

19 ☐ Module 4: Gathering Evidence and Speaking to Others: Natural Disasters in the Western Hemisphere

Sixth Grade ELA

20 ☐ Module 1: Reading Closely and Writing to Learn: Myths: Not Just Long Ago

21 ☐ Module 2A: Working with Evidence: Rules to Live By

22 ☐ Module 2B: Voices of Adversity

23 ☐ Module 3A: Perspective: The Land of the Golden Mountain

24 ☐ Module 3B: Reading Closely and Writing to Learn: Point of View and Perspective

25 ☐ Module 4: Reading for Research and Writing an Argument: Insecticides: Costs vs. Benefits

Seventh Grade ELA

26 ☐ Module 1: Reading Closely and Writing to Learn: Journeys and Survival

27 ☐ Module 2A: Working with Evidence: Working Conditions

28☐ Module 2B: Working with Evidence Identity and Transformation: Then and Now

29 ☐ Module 3: Understanding Perspectives: Slavery: The People Could Fly

30 ☐ Module 4A: Reading and Research: Screen Time and the Developing Brain

31 ☐ Module 4B: Reading and Research: Water is Life

Eighth Grade ELA

32 ☐ Module 1: Reading Closely and Writing to Learn: Finding Home: Refugees

33 ☐ Module 2A: Working with Evidence: Taking a Stand

34 ☐ Module 2B: Working with Evidence: A Midsummer Night's Dream and the Comedy of Control

35 ☐ Module 3A: Understanding Perspectives: Japanese-American Relations in WWII

36 ☐ Module 3B: Understanding Perspectives: The Civil Rights Movement and the Little Rock Nine

37 ☐ Module 4: Research, Decision-Making, and Forming Positions: Sustainability of the Worlds' Food Supply

E15. To what degree has participating in the Teacher Potential Project influenced your overall approach to teaching?

Select one only

- ❍ Not at all..... 1
- ❍ Somewhat..... 2
- ❍ A good deal..... 3
- ❍ Extensively..... 4
- ❍ Not Applicable..... 99
- ❍ NO RESPONSE ..... M

E16. To what degree has participating in the Teacher Potential Project Influenced your overall approach to your understanding of your state standards?

Select one only

- ❍ Not at all..... 1
- ❍ Somewhat..... 2
- ❍ A good deal..... 3
- ❍ Extensively..... 4
- ❍ Not Applicable..... 99
- ❍ NO RESPONSE ..... M

E17. To what degree has participating in the Teacher Potential Project influenced your overall approach to the ways you teach English Language Arts (ELA)?

Select one only

- ❍ Not at all..... 1
- ❍ Somewhat..... 2
- ❍ A good deal..... 3
- ❍ Extensively..... 4
- ❍ Not Applicable..... 99
- ❍ NO RESPONSE ..... M

E18. To what degree has participating in the Teacher Potential Project influenced your overall approach to the ways you teach other subjects?

Select one only

- ❍ Not at all 1
- ❍ Somewhat 2
- ❍ A good deal 3
- ❍ Extensively 4
- ❍ Not Applicable 99
- ❍ NO RESPONSE ..... M

E19. To what degree has participating in the Teacher Potential Project influenced your overall approach to the ways you manage your classroom?

Select one only

- ❍ Not at all 1
- ❍ Somewhat 2
- ❍ A good deal 3
- ❍ Extensively 4
- ❍ Not Applicable 99
- ❍ NO RESPONSE ..... M

E20. To what degree has participating in the Teacher Potential Project influenced your overall approach to the ways you engage students in rigorous learning tasks?

Select one only

- ❍ Not at all 1
- ❍ Somewhat 2
- ❍ A good deal 3
- ❍ Extensively 4
- ❍ Not Applicable 99
- ❍ NO RESPONSE ..... M

F. FUTURE TEACHING PLANS

This section collects information about your future teaching plans.

F1. How long do you plan to remain in the K-12 teacher position?

Select one only

- ❍ As long as I am able 1
- ❍ Until I am eligible for retirement and/or Social Security benetits 2
- ❍ Until a specific life event occurs (e.g., parenthood, marriage) 3
- ❍ Until a more desirable job opportunity comes along 4
- ❍ I plan to leave as soon as I can 5
- ❍ Undecided at this time 6

F2. Approximately how many years do you think you will remain in teaching after this school year?

|___|___| YEARS

F3. IF YOUR SCHOOL IS PARTICIPATING IN EL EDUCATION'S TEACHER POTENTIAL PROJECT: **Will the EL Education TPP program affect <u>whether you decide to stay in teaching</u> in the next school year?**

- ❍ Yes 1
- ❍ Possibly 2
- ❍ No 0

F4. IF YOUR SCHOOL IS PARTICIPATING IN EL EDUCATION'S TEACHER POTENTIAL PROJECT: **Will the EL Education TPP program affect <u>where you decide to teach</u> in the next school year?**

- ❍ Yes 1
- ❍ Possibly 2
- ❍ No 0

G. BACKGROUND INFORMATION

G1a1. Do you have a bachelor's degree?

- ❍ Yes ..... 1 GO TO G1a2
- ❍ No ..... 0 GO TO G1b1

NO RESPONSE ..... M GO TO G1b1

G1a2. What field of study do you have your bachelor's degree in?

[ ] GO TO G1a3

NO RESPONSE ..... M GO TO G1a3

G1a3. In what year was your bachelor's degree awarded?

[ ]

NO RESPONSE ..... M

G1b1. Do you have a master's degree?

- ❍ Yes ..... 1 GO TO G1b2
- ❍ No ..... 0 GO TO G1c1

NO RESPONSE ..... M GO TO G1c1

G1b2. What field of study do you have your master's degree in?

[ ] GO TO G1b3

NO RESPONSE ..... M GO TO G1b3

G1b3. In what year was your master's degree awarded?

[ ]

NO RESPONSE ..... M

G1c1. Do you have an education specialist or professional diploma based on at least one year's work past master's degree?

- ❍ Yes ..... 1 GO TO G1c2
- ❍ No ..... 0 GO TO G1d1

NO RESPONSE ..... M GO TO G1d1

G1c2. What field of study do you have your <u>education specialist or professional diploma based on at least one year's work past master's degree</u> in?

[ ] GO TO G1c3

NO RESPONSE ..... M GO TO G1c3

G1c3. In what year was your <u>education specialist or professional diploma based on at least one year's work past master's degree</u> awarded?

[ ]

NO RESPONSE ..... M

G1d1. Do you have a <u>doctoral</u> degree?

- ❍ Yes ..... 1 GO TO G1d2
- ❍ No ..... 0 GO TO G1e1

NO RESPONSE ..... M GO TO G1e1

G1d2. What field of study do you have your <u>doctoral</u> degree in?

[ ] GO TO G1d3

NO RESPONSE ..... M GO TO G1d3

G1d3. In what year was your <u>doctoral</u> degree awarded?

[ ]

NO RESPONSE ..... M

G1e1. Do you have a <u>professional degree (e.g., M.D., L.L.B., J.D., D.D.S.)</u>?

- ❍ Yes ..... 1 GO TO G1e2
- ❍ No ..... 0 GO TO G1f1

NO RESPONSE ..... M GO TO G1f1

G1e2. What field of study do you have your <u>professional degree (e.g., M.D., L.L.B., J.D., D.D.S.)</u> in?

[ ] GO TO G1e3

NO RESPONSE ..... M GO TO G1e3

G1e3. In what year was your professional degree (e.g., M.D., L.L.B., J.D., D.D.S.) awarded?

NO RESPONSE ..... M

G2. Which of the following best describes the teaching certificate you currently hold in the state in which you currently teach?

Select one only

- ❍ Regular or standard state certificate or advanced professional certificate..... 1
- ❍ Probationary certificate (the initial certificate issued after satisfying all requirements except the completion of a probationary period) 2
- ❍ Provisional or other type of certificate given to persons who are still participateing in what the state calls an "alternative certification program"..... 3
- ❍ Temporary certificate (requires some additional college coursework and/or student teaching before regular certification can be obtained) 4
- ❍ Emergency certificate or waiver (issued to persons with insufficient teacher preparation who must complete a regular certification program in order to continue teaching) 5
- ❍ I do not hold any of the above certifications in THIS state ..... 6 ➔ GO TO G3a

G3. In what area(s) does the [FILL teaching certificate marked in G2] allow you to teach in your state?

SELECT ALL THAT APPLY

- ❍ Preschool (birth-PreK) 1
- ❍ Elementary (K-5)..... 2
- ❍ Middle grades (6-8) 3
- ❍ Secondary grades (9-12)..... 4
- ❍ Specific subject areas (K-12) *(specify)* 5

  Specify

- ❍ Other *(specify)* 6

  Specify

G3a. What subject(s) are you currently teaching at your school?

GO TO G3b

G3b. What grade level(s) are you currently teaching at your school?

GO TO G4

G4. Do you hold a valid regular or standard certification from a state other than the one in which you are currently teaching?

- ❍ Yes 1
- ❍ No 0

NO RESPONSE ..... M

G5. Have you enrolled in college or university courses inREFERENCE YEAR?

- ❍ Yes 1
- ❍ No ..... 0 GO TO G9

NO RESPONSE ..... M GO TO G9

G6. Which of the following best describes your enrollment in these courses?

Select one only

- ❍ Individual courses (not pare of a program leading to a degree or certificate) 1
- ❍ Vocational certificate program 2
- ❍ Associate degree granting program 3
- ❍ Bachelor's degree granting program 4
- ❍ Master's degree granting program 5
- ❍ Education specialist or professional diploma program (at least one year beyond master's level) 6
- ❍ Certificate of Advanced Graduates Studies program 7
- ❍ Doctorare or professional degree granting program (Ph.D., Ed.D., M.D., J.D., D.D.S.) 8

G7. Which of the following best describes the reason you enrolled in these courses?

Select one only

- ❍ To obtain for use in a K-12 teaching position 1
- ❍ To obtain for use in a position in the field of education but not as a K-12 teacher 2
- ❍ To obtain for use in a position outside the field of education 3
- ❍ For reasons unrelated to obtaining or using in a job (e.g. personal fulfillment) 4

G8. Were these courses needed to obtain, renew, or maintain teaching certification?

- ❍ Yes 1
- ❍ No 0

G9. Including the current school year, how long have you worked as a teacher in the following settings?

Please do not include student or substitute teaching. IF NONE, PLEASE ENTER: 00

ENTER ONE RESPONSE PER ROW;
IF NONE, PLEASE ENTER: **00**

	YEARS	MONTHS
a. Total amount of time teaching in current school	|___|___|	|___|___|
b. Total amount of time teaching in current district	|___|___|	|___|___|
c. Total amount of time teaching in any school	|___|___|	|___|___|

G9a Including the current school year, have you worked as a substitute teacher?

- ❍ Yes 1
- ❍ No 0 G10
- ❍ Don't Know ..... D G10

NO RESPONSE ..... M G10

G9b Including the current school year, how long have you worked as a substitute teacher in the following settings?

Please do not include student teaching. IF NONE, PLEASE ENTER: 00

ENTER ONE RESPONSE PER ROW;
IF NONE, PLEASE ENTER: 00

	YEARS	MONTHS
a. Total amount of time in part-time substitute teaching assignments	|___|___|	|___|___|
b. Total amount of time in full-time substitute teaching assignments	|___|___|	|___|___|

G10. In the last 12 months, have you applied for a job in an attempt to leave the position of a pre-K-12 teacher?

- ❍ Yes 1
- ❍ No 0

G11. What is your gender?

- ❍ Male 1
- ❍ Female 2
- ❍ Other *(specify)* 3

Specify

G12. Are you Hispanic or Latino?

Select one only

- ❍ No, not of Hispanic, Latino, or Spanish origin 1
- ❍ Yes, of Hispanic, Latino, or Spanish origin 2

G13. What is your race?

SELECT ALL THAT APPLY

- ❍ White 1
- ❍ Black or African American 2
- ❍ American Indian or Alaskan Native 3
- ❍ Asian or Pacific Islander 4
- ❍ Some other race *(specify)* 5

Specify

1

G14. What is your age?

|___|___| AGE

Appendix D:

Teacher observation instrument

This page has been left blank for double-sided copying.

MATHEMATICA Policy Research

Expeditionary Learning Evaluation
Investing in Innovation Validation

INSTRUCTION: PLEASE MARK ONE RESPONSE FOR EACH ITEM UNLESS OTHERWISE NOTED. MARK (X) BOXES NEATLY

Observer Name:_____

OBSERVER ID: |__|__|__|__|__|__|__|__|

TEACHER NAME:_____

TEACHER MPRID: : |__|__|__|__|__|__|__|__|__|

NUMBER OF TEACHERS: |__|__|__|

NUMBER OF STUDENTS: |__|__|__|

GRADE: |__|__|__|

SCHOOL:_____

☐ CONTROL ☐ TREATMENT

☐ NOVICE ☐ EXPERIENCED

EL CURRICULUM: _____

COMMENTS (PLEASE DESCRIBE THE LESSON AND ANY IMPORTANT INFORMATION REGARDING THE CLASS PERIOD).

Date	|__|__| / |__|__| / **20** |__|__| Month Day Year
Day of the Week:	1 ☐ Monday 2 ☐ Tuesday 3 ☐ Wednesday 4 ☐ Thursday 5 ☐ Friday

Enter Observation Start Time and End Time:

		Mark AM OR PM
Start:	|__|__| : |__|__|	1 ☐ AM 2 ☐ PM
End:	|__|__| : |__|__|	1 ☐ AM 2 ☐ PM

Which of the following reflects your presence in the class?

1 ☐ Missed the beginning of the class period

2 ☐ Missed the ending of the class period

3 ☐ Present for the entire class period

SECTION 1: GENERAL INSTRUCTION

A. LESSON STRUCTURE

1. Did the teacher and/or students explicitly state the purpose of instruction out loud? (i.e., Learning Target or Goal, "I Can" statement, Students Will Be Able To…) ***2. If yes, please write or paraphrase the Learning Target/Goal below:***	1 ☐ Yes 0 ☐ No
3. During the lesson, did the teacher refer back to the purpose/learning goals? (i.e., restating or pointing to purpose on the board) **Note: This does not include if the teacher stated the purpose or learning goals at the end of the class.**	1 ☐ Yes 0 ☐ No
4. Did the teacher call attention to students' prior knowledge, such as things they learned in other subjects or earlier lessons or their personal experiences? (i.e., connecting lesson or topic to students' personal experience or things learned in school.)	1 ☐ Yes 0 ☐ No
5. Did the teacher make connections to the real world? (i.e., usefulness or relevance of lesson or topic to real world contexts beyond the classroom)	1 ☐ Yes 0 ☐ No
6. Did the teacher appear to be well organized and prepared to conduct the lesson? (i.e., teacher had a written plan/notes, presented the lesson concisely, had any necessary materials ready, was sure of what to do/say, lesson fit timeframe, had prepared instructions, etc.)	1 ☐ Yes 0 ☐ No
7. Did the lesson include a summary at the end, to synthesize learning?	1 ☐ Yes, but only the teacher summarized the lesson. 2 ☐ Yes, teacher and the students summarized the lesson. 0 ☐ No, neither the teacher nor the students summarized the lesson.
8. Which instructional groupings were used during the class period? **MARK ALL THAT YOU OBSERVE**	1 ☐ Whole class 2 ☐ Independent work 3 ☐ Pair/partner work 4 ☐ Small group work without the teacher 5 ☐ Small group work with the teacher 6 ☐ Other (specify)_____
9. How many times did the teacher switch instructional groupings? **Note: This is how many times the teacher switched instructional groupings, even if the same types of groupings are used (for example, whole class to independent to whole class would be 2 switches).**	0 ☐ None, the teacher used the same instructional grouping throughout the whole class period 1 ☐ 1 2 ☐ 2 3 ☐ 3

	4 ☐ 4 or more
10. How would you rate the overall coherence of the lesson? (Were concepts and activities linked to one another within the lesson?)	N ☐ NA – only one activity took place during the class period. 1 ☐ Low (Lesson was disjointed; activities didn't seem to relate to each other.) 2 ☐ Medium (Most of lesson was coherent; one or two activities may not have related to the others.) 3 ☐ High (All of the lesson was coherent; beginning, middle, end activities clearly related to each other.)

B. CLASSROOM DISCUSSION

11. Did classroom discussion occur?	1 ☐ Yes 0 ☐ No
12. Which of the following is most characteristic of who spoke most in classroom?	N ☐ NA - No discussion occurred. 1 ☐ The teacher spent the most time talking. 2 ☐ The teacher and the students spent an equal amount of time talking. 3 ☐ The students spent the most time talking.
13. Which of the following best describes how the teacher encouraged students to participate in the discussion? **MARK ALL THAT YOU OBSERVE**	N ☐ NA -No discussion occurred. 1 ☐ The teacher called on the same 2-3 students. 2 ☐ The teacher called on a range of students (more than the same 2-3). 3 ☐ The teacher called on a range of students (more than the same 2-3) AND had a system to ensure s/he called on a range of students (e.g., the teacher calls on students in alphabetical order or calls on students in order of desk rows). 4 ☐ Teacher poses questions generally to the whole class. 5 ☐ Teacher cold-called students.
14. Which of the following is most characteristic of how the teacher incorporated student input into the lesson? **MARK ALL THAT YOU OBSERVE**	N ☐ NA -No discussion occurred. 1 ☐ The teacher never or rarely incorporated student questions, comments, and ideas into the lesson. 2 ☐ The teacher incorporated student questions, comments, and ideas into the lesson verbally (i.e.: repeated a student's question, answer, or comment). 3 ☐ The teacher incorporated student questions, comments, and ideas into the lesson in writing (i.e.: put a student's work or questions on the board or on an overhead).

15. Which of the following is most characteristic of participation in the classroom? (*Participation is defined as students raising their hands, asking and/or answering questions, and/or actively discussing with peers.)*	N ☐ NA - No discussion occurred 1 ☐ Less than half of the class participated 2 ☐ At least half of the class participated, but students didn't ask their own questions 3 ☐ At least half of the class participated and students asked their own questions
16. What types of feedback did the teacher give the students when the teacher asked questions? **MARK ALL THAT YOU OBSERVE** (***Higher-level thinking*** *requires that the students engage in analysis, synthesis, comparing/contrasting, cause/effect, critiquing, or evaluating. Any thinking above the level of reporting, copying or describing is considered higher-level.)*	N ☐ NA – No discussion occurred 1 ☐ The teacher gave non-verbal feedback (such as nodding or shaking their head). 2 ☐ The teacher let the student know if their response was right or wrong. 3 ☐ The teacher asked the student probing questions, for more information, or follow-up questions. 4 ☐ The teacher encouraged the student to use more complex language. 5 ☐ The teacher encouraged the student to think at a higher level (see definition). 6 ☐ The teacher asked other students to help answer the question.
16a. Which type of feedback did you see most often?	N ☐ NA – No discussion occurred 1 ☐ The teacher gave non-verbal feedback (such as nodding or shaking their head). 2 ☐ The teacher let the student know if their response was right or wrong. 3 ☐ The teacher asked the student probing questions, for more information, or follow-up questions. 4 ☐ The teacher encouraged the student to use more complex language. 5 ☐ The teacher encouraged the student to engage in higher-level thinking. 6 ☐ The teacher asked other students to help answer the question.
16b. Did the teacher provide mostly positive or mostly negative feedback to the students?	N ☐ NA – No discussion occurred 1 ☐ Mostly positive feedback 2 ☐ Mostly negative feedback 3 ☐ Equally positive and negative feedback
17. How would you rate the thoughtfulness of the classroom discussion?	N ☐ NA -No discussion occurred 1 ☐ Low (Students' responses were automatic or rote.) 2 ☐ Medium (Student(s) sometimes responded thoughtfully.) 3 ☐ High (Student(s) sometimes responded thoughtfully plus generated their own probing questions.)
17a. How often did students respond to peer comments and/or build on each other's thinking?	N ☐ NA -No discussion occurred 1 ☐ Never or rarely 2 ☐ Sometimes 3 ☐ Frequently

C. CRITICAL THINKING

18. What kinds of questions did the teacher ask during the class period either verbally or in writing? (e.g., questions asked in class assignments) **MARK ALL THAT YOU OBSERVE** *MAKE PREDICTIONS: Based on what you've read so far, what do you think will happen next?* *MAKE INFERENCES: Based on what the character did, how do you think she felt?* *ANALYZE: Classify, compare/contrast, cause/effect, problem/solution, critique/evaluation* *SYNTHESIZE: Draw conclusions (What is the moral of this story? What is the main point the writer is trying to make?)*	The teacher asked: 0 ☐ The teacher did not ask any questions. 1 ☐ Factual, yes/no, basic questions. 2 ☐ Students to explain 'why' or 'how' they arrived at an answer, or a 'why'/'how' question more generally. 3 ☐ Students to make predictions (see definition). 4 ☐ Students to make inferences, analyze, synthesize, draw conclusions, analyze, evaluate, and/or critique (see definitions). 5 ☐ Students to cite evidence from the text to support their responses. 6 ☐ Students to use a discussion protocol to guide class discussion.
19. Which of the following is most characteristic of the questions the teacher asked either verbally or in writing? *Basic questions require students to record or report on basic information or facts.* *Higher-order-thinking (HOT) questions require that the students engage in analysis, synthesis, comparing/contrasting, cause/effect, critiquing, or evaluating.*	N ☐ NA. The teacher did not ask any questions. 0 ☐ Unable to tell. 1 ☐ The teacher asked mainly basic, factual or yes/no questions. 2 ☐ The teacher asked an even mix of basic and higher-order-thinking (HOT) questions. 3 ☐ The teacher asked mainly higher-order-thinking (HOT) questions.
20. In general, what was the level of thinking reinforced during the lesson? *Basic questions require students to record or report on basic information or facts.* *Higher-order-thinking (HOT) questions require that the students engage in analysis, synthesis, comparing/contrasting, cause/effect, critiquing, or evaluating.*	N ☐ NA. There was no classroom discussion. 0 ☐ Unable to tell. 1 ☐ Low (Students were mostly engaged in basic, factual, yes/no thinking.) 2 ☐ Medium (Students sometimes engaged in HOT.) 3 ☐ High (Students often engaged in HOT.)

D. GENERAL INSTRUCTIONAL FORMATS OR TECHNIQUES

21. Which instructional formats or techniques did the teacher use during the class period? **MARK ALL THAT YOU OBSERVE**	1 ☐ Used chalk board, smart board, projector, or chart paper 2 ☐ Extended teacher talk/ lecture (little to no student involvement) 3 ☐ Q & A (teacher initiates back and forth with students, mostly single-turn) 4 ☐ Discussion protocol (e.g., Socratic seminar, Think-Pair-Share) 5 ☐ Project-based (extended activity with presentation or product) 6 ☐ Games/role plays/drama 7 ☐ Multimedia (film/videos/music/art) 8 ☐ Computers (students or teacher using computers) 9 ☐ Text-based protocol (e.g., Notecatcher) 10 ☐ Group/pair work 11 ☐ Other (specify) _____
22. How would you rate the pacing of instructional formats or techniques used during the class period?	1 ☐ Too fast (the pace of activities was too hectic, distracting) 2 ☐ The right pace (kept most students' attention on task) 3 ☐ Too slow (most students' attention wandered during long activities)
23. What strategies did the teacher use to scaffold instructional activities to help students understand complex material and/or content? **MARK ALL THAT YOU OBSERVE**	0 ☐ The teacher did **not** scaffold instruction in any form. 1 ☐ Breaking ideas/questions into smaller or simpler segments 2 ☐ Re-representing information/ideas visually 3 ☐ Having students explain ideas to each other 4 ☐ Using examples, having students generate examples, and/or referring to real world examples 5 ☐ Returning to students who seem to struggle, to see if they now understand

	6 ☐ Providing more structured activities or tasks for some students 7 ☐ Round Robin reading or other forms of students reading aloud 8 ☐ Teachers reading for students

E. ASSESSMENT

24. During the class period, did the teacher give the students a quiz or test?	1 ☐ Yes 0 ☐ No
25. During the class period, did the teacher or students use a rubric to assess student work?	1 ☐ Yes 0 ☐ No
26. During the class period, what types of informal assessment did you observe? These can be verbal or non-verbal and individual or classroom-wide. **MARK ALL THAT YOU OBSERVE.**	0 ☐ The teacher did not use any types of informal assessment. 1 ☐ The teacher walks around the classroom, observing students during independent, paired work or group work. 2 ☐ The teacher walks around the classroom, marking students' work. 3 ☐ The teacher uses exit tickets. 4 ☐ The teacher uses quick "MARKS". 5 ☐ The teacher asks for a show of hands, thumbs up/down, red light/green light to assess students' understanding of the lesson. 6 ☐ Other (specify) _____
28. During the class period, did students provide verbal or written feedback to other students? *This includes feedback provided as part of peer conferencing (peer editing or peer critique), as well as less formal feedback, such as silent agreement.*	1 ☐ Yes 0 ☐ No

SECTION 1 NOTES:

SECTION 2: READING AND WRITING INSTRUCTION

A. BUILDING ACADEMIC VOCABULARY AND CONTENT KNOWLEDGE

29. Did the teacher or students use <u>academic vocabulary</u> during the class period? *Academic Vocabulary: Content or subject specific terms or complex or challenging words. This depends on grade level and includes words that are infrequently used, multi-syllabic, or about complex concepts.*		1 ☐ Yes 0 ☐ No
30. What strategies did the teacher use to provide instruction on <u>academic vocabulary</u>? **MARK ALL THAT APPLY**	N ☐ <u>NA</u> – no instruction on academic vocabulary 1 ☐ Gave the most common "dictionary" definition 2 ☐ Used context clues to infer word meaning 3 ☐ Used a real world or personal example to define a word (making connections to background knowledge) 4 ☐ Offered synonyms/antonyms	5 ☐ Used the word in a sentence 6 ☐ Used visuals or gestures 7 ☐ Asked students for definition 8 ☐ Used text to develop students' vocabulary 9 ☐ Other:
31. To what degree did the teacher focus on students' <u>academic vocabulary</u> instruction?	N ☐ <u>NA</u> – no instruction on academic vocabulary 1 ☐ <u>Minimally</u>. The teacher mentioned or called attention to academic vocabulary briefly or in passing. 2 ☐ <u>Somewhat</u>. The teacher focused on academic vocabulary during part of the class period (e.g., mini-lesson for 5-10 minutes). 3 ☐ <u>Very</u>. The teacher focused on academic vocabulary for almost the entire or the entire class period.	
32. Did the lesson involve developing students' <u>content knowledge</u>? *Content Knowledge: Includes information from the content subjects. This includes facts, major ideas from these subjects, information that develops students' understanding of the physical world (science and physical education), the human world (social studies, arts, philosophy) and tools of human thought (mathematics). E.g., historic context of a novel*		1 ☐ Yes 0 ☐ No
33. To what degree did the teacher engage students in building their <u>content knowledge</u>? *Engaged students: students answering questions, discussing, writing, acting, etc.*	N ☐ <u>NA</u> – content knowledge was not used during the class period 1 ☐ <u>Very little or not at all</u>. The teacher did all of the work in the transfer of knowledge. 2 ☐ <u>Somewhat</u>. Students did some of the work, but the teacher still did most of the work. Students responded to questions but the teacher did most of the transfer of knowledge. 3 ☐ <u>A great deal</u>. Students did most of the work. Students located the knowledge or became owners of the knowledge.	
34. How much of the class period did the teacher spend on developing students' <u>content knowledge</u>?	N ☐ <u>NA</u> – content knowledge was not used during the class period 1 ☐ <u>Minimally.</u> The teacher mentioned or called attention to content knowledge briefly or in passing.	

	2 ☐ Somewhat. The teacher focused on content knowledge during part of the class period (i.e. mini-lesson on a content topic) 3 ☐ Very. The teacher focused on content knowledge for almost the entire or the entire class period.
35. What was the overall quality of content knowledge instruction?	N ☐ NA – content knowledge was not used during the class period 1 ☐ Low (Content knowledge was not memorable; it may have been confusing and/or disconnected to the lesson.) 2 ☐ Medium (Content knowledge was somewhat memorable, but may not have been clearly connected to the main point of the lesson.) 3 ☐ High (Content knowledge was very memorable and clearly connected to the main point of the lesson.)

B. READING ACTIVITIES

36. What kind of reading activities did the students engage in during the class period? (with or without the teachers' help) **MARK ALL THAT APPLY**	N ☐ The students did not engage in reading activities during the class period. 2 ☐ Pre-reading (e.g. introducing the topic before reading, vocabulary review, calling attention to text features) 3 ☐ Making predictions 4 ☐ Connecting text to life 5 ☐ Looking up new words 6 ☐ Determining word meanings from text 7 ☐ Annotating text 8 ☐ Close reading (students re-read text to look for answers to specific questions) 9 ☐ Analyzing Text Structure 10 ☐ Summarizing (Re-telling, Paraphrasing, "Gist")	11 ☐ Students reading aloud (e.g., round robin) 12 ☐ Teacher reading aloud 13 ☐ Analyzing text for themes, plot, and/or character motives 14 ☐ Reviewing text to find evidence to support a response or viewpoint 15 ☐ Analyzing models to develop criteria for student work (i.e.: guidelines) 16 ☐ Students silently reading a text of choice 17 ☐ Students silently reading an assigned text. 18 ☐ Other (specify) _____
36a. What type of pre-reading occurred? **MARK ALL THAT APPLY**	N ☐ The students did not engage in reading during the class period. 0 ☐ No pre-reading occurred. 1 ☐ The teacher provided information to prepare students to read (e.g., reviewing vocabulary in the text, explaining the purpose of reading). 2 ☐ The teacher summarized the text in detail so that students did not need to read the text to answer post-reading questions.	

37. How many texts were used in the lesson during the class period? **37a. Record the title of text, if known.**	N ☐ NA The students did not engage in reading during the class period. |___|___| NUMBER OF TEXTS	*RECORD TEXT TITLES BELOW:* a. _____ b. _____ c. _____ d. _____

37b. Mark the text type or genre of the texts used during the class period.	N ☐ NA – No reading activities during class period. 0 ☐ Text type/genre unknown 1 ☐ Novel 2 ☐ Novella 3 ☐ Fictional short stories 4 ☐ Poetry 5 ☐ Drama/plays 6 ☐ Biography, autobiography, memoir 7 ☐ Personal essays 8 ☐ Journal, diary	9 ☐ Nonfictional short stories 10 ☐ Speeches 11 ☐ Magazine, newspaper, internet, or journal articles 12 ☐ Textbooks, technical writings 13 ☐ Reference books (dictionary, thesaurus, encyclopedia, atlas) 14 ☐ Curricula, resources, or eBooks from the web 15 ☐ Photos, videos, or media clips 16 ☐ Other (specify) _____
38. Did the teacher ask students to reread the text? ("Can you show me where in the text you found that?")	N ☐ NA – No reading activities during class period. 1 ☐ Yes 0 ☐ No	
38a. Did students reread the text without the teacher's prompting?	N ☐ NA – No reading activities during class period. 1 ☐ Yes 0 ☐ No	
39. Were the reading activities focused on the meaning of the text? (e.g., main idea, key information, main theme, characters, plot)		N ☐ NA – No reading activities during class period. 1 ☐ Not very much

	2 ☐ Somewhat 3 ☐ Very much
40. How much class time did the students spend reading independently?	N ☐ NA – No reading activities during class period. 1 ☐ Barely any time 2 ☐ Less than half of the period 3 ☐ About half of the period 4 ☐ More than half of the period
41. Did the students have the opportunity to read a "book of choice" during the class period? (Such as during Drop Everything and Read (DEAR), Sustained Silent Reading (SSR), free read time.)	1 ☐ Yes 0 ☐ No

C. WRITING ACTIVITIES

42. What kinds of writing activities did the students do during the class period (with or without the teacher's help)? **MARK ALL THAT APPLY**	0 ☐ The students did not engage in writing activities during the period. 1 ☐ Pre-writing as part of the writing process (e.g. brainstorming before writing) 2 ☐ Creating an outline or graphic organizer 3 ☐ Drafting as part of the writing process 4 ☐ Revising/editing methods as part of the writing process 5 ☐ Peer Conferencing (Peer Critique Sessions, Peer Editing) as part of the writing process 6 ☐ Using a graphic organizer or worksheet provided by the teacher 7 ☐ Short responses to questions that promote thinking (such as "think-tasks"), warm up thinking 8 ☐ On-demand writing prompts (>2 paragraphs)	9 ☐ Citing sources or evidence 10 ☐ Vocabulary development 11 ☐ Writing conventions (grammar, mechanics, spelling, paragraphing, formatting [such as use of subheadings, use of quotations]) 12 ☐ Text structures (e.g., poetry structure [stanzas], essay structure) 13 ☐ Quiz or exam 14 ☐ Notetaking, copying and/or recording information annotating text 15 ☐ Analyzing models to develop criteria for student work (i.e., protocol, rubric) 16 ☐ Other (specify) _____
43. What was the genre or mode of the writing or writing activities that the students engaged in? **MARK ALL THAT APPLY**	N ☐ NA – No writing activities during class period. Fiction: 1 ☐ Poem 2 ☐ Short Story	Non-fiction: 3 ☐ Biography, Memoir, Personal Essay 4 ☐ Narrative, How To 5 ☐ Descriptive 6 ☐ Informational/Expository 7 ☐ Opinion writing 8 ☐ Literary Analysis 9 ☐ Book reviews or literature reviews	10 ☐ Persuasion (No sources, references) 11 ☐ Argument writing (Using sources, references) 12 ☐ Free writing (journaling, brainstorming) 13 ☐ Worksheet, graphic organizer, notetaking, or exam 14 ☐ Letters 15 ☐ Speeches 16 ☐ Other (specify) _____ 17 ☐ Couldn't tell/Unknown
44. What is the purpose of the writing? **MARK ALL THAT APPLY**	N ☐ NA – No writing activities during class period. 1 ☐ Practicing grammar, mechanics, and/or spelling 2 ☐ Literary (poetry, personal essays, fictional stories) 3 ☐ Making arguments based on evidence 4 ☐ Writing to demonstrate understanding of text 5 ☐ Promote personal reflection on the text	8 ☐ Recording information/ideas 9 ☐ Writing from sources (using text or notes in their writing) 10 ☐ Literary analysis (analyzing text for themes, plot, and/or character motives 11 ☐ Analyze text structure 12 ☐ Other (specify) _____	

	6 ☐ Synthesize what students know 7 ☐ Learning of content knowledge	13 ☐ Couldn't tell
45. How long was the writing? **MARK ALL THAT APPLY**	N ☐ NA – No writing activities during class period. 1 ☐ Words or lists of words and phrases, bulleted points 2 ☐ Short answer: 1-2 sentences 3 ☐ Long answer/paragraph: 3-6 sentences 4 ☐ Short: 2-4 paragraphs	5 ☐ Long: 5-8 paragraphs 6 ☐ Reports: more than 8 paragraphs 7 ☐ Other (specify) _____ 8 ☐ Couldn't tell
46. How much of the class period did the students spend on this writing activity?	N ☐ NA – No writing activities during class period. 1 ☐ Barely any time 2 ☐ Up to half of the period 3 ☐ More than half of the period	

SECTION 2 NOTES:

Section 3: Classroom Management and Environment

A. CLASSROOM MANAGEMENT/MANAGING STUDENT BEHAVIOR/RESOURCES

47. Were behavioral rules, norms, or expectations clearly posted somewhere in the room?	1 ☐ Yes 0 ☐ No
48. How frequently did students disrupt the class? *Disruption is defined as behavior that takes away instructional time from at least one other student (other than the student who is disruptive). So, a student sitting at his or her desk scribbling on paper might be disruptive, but unless it's taking away learning time from another student (e.g., student at desk next to disruptive student is distracted by the scribbling and not paying attention to the lesson), it should not be coded.*	N ☐ There was no student disruption 1 ☐ Rarely 2 ☐ Somewhat Frequently (a few times) 3 ☐ Very Frequently
49. Which of the following best describes how teachers enforced behavioral rules in the classroom?	N ☐ N/A – Behavioral rules were not demonstrated because there was no student disruption. 1 ☐ Behavioral rules and/or expectations were not enforced. 2 ☐ Behavioral rules and/or expectations were inconsistently enforced. 3 ☐ Behavioral rules and/or expectations for behavior were consistently enforced.
50. How effective were the teacher's attempts to re-direct student disruptions?	N ☐ N/A – There was no student disruption. 1 ☐ The teacher did not attempt to redirect student disruption. 2 ☐ The teacher's attempts to redirect student disruption are mostly ineffective. 3 ☐ The teacher's attempts to redirect student disruption are mostly effective.
51. Did the teacher's redirection of student disruption(s) take away from learning time? *Examples of redirection include using proximity, hand signals, whispering to student, writing names on the board.*	N ☐ N/A – there was no student disruption. 1 ☐ The teacher did not attempt to redirect student disruption. 2 ☐ Yes, it took away from learning time (at least 10 minutes total during the class period). 3 ☐ No, it did not take away from learning time (less than 10 minutes total during class period).
52. Did the teacher use punitive control to manage student disruption? *Examples include yelling, threats, punishment, public shaming, and disrespectful language.*	1 ☐ Yes 0 ☐ No or NA
52a. Did the teacher use positive reinforcement to manage student behavior? *Examples include giving students merits, compliments, and rewards for good behavior.*	1 ☐ Yes 0 ☐ No
53. Was instructional time (10 minutes or more) lost because of wait time or transitions?	1 ☐ Yes

	0 ☐ No or NA
54. Was the classroom well-organized and neat?	1 ☐ Yes 0 ☐ No
55. Were materials easily accessible to students? (Including books, pencils, project items, computers, etc.)	1 ☐ Yes 0 ☐ No
56. Was the classroom set up to encourage and facilitate peer collaboration (are the tables set up in pairs or groups, etc.)?	1 ☐ Yes 0 ☐ No
57. Was student work posted in the classroom?	1 ☐ Yes 0 ☐ No
58. Did posted student work include a posted standard that links work to academic content or topics? (E.g., student-made posters about a novel read in class)	N ☐ N/A – There was no student work posted. 1 ☐ Yes 0 ☐ No
59. Was there a library of books in the classroom (not including textbooks, dictionary, and thesauruses)?	1 ☐ Yes 0 ☐ No
59a. Did the classroom have a shelf of textbooks, dictionaries, and thesauruses that were accessible to students?	1 ☐ Yes 0 ☐ No
59b. Were there any computers or tablets in the classroom that were accessible to students (e.g., a workstation)?	1 ☐ Yes 0 ☐ No

B. CLASSROOM CULTURE

How characteristic are the following descriptions of the classroom?

60. The teacher and students demonstrated respect for one another. *(Demonstrates respect: using eye contact, warm voices, respectful language, and cooperation, and/or sharing)*	1 ☐ Not very 2 ☐ Somewhat 3 ☐ Very
61. The teacher was warm, empathetic, and/or supportive.	1 ☐ Not very 2 ☐ Somewhat

	3 ☐ Very
62. Classroom climate was characterized by inclusion of all. (*Inclusion does not refer to classes containing both special education and general education students.)*	1 ☐ Not very 2 ☐ Somewhat 3 ☐ Very
63. The teacher successfully encouraged students to interact with each other (e.g. teacher prompted students to work collaboratively in pairs or groups, discussion in these pairs/groups was relevant to the content of the lesson, teacher prompted students to attend to peers' comments).	1 ☐ Not very 2 ☐ Somewhat 3 ☐ Very
64. How many students were receptive to the teacher's instructions and feedback? (e.g. students' eyes were on the teacher when the teacher was speaking; students showed their understanding with nodding, raised their hands to answer questions)	1 ☐ None or few students 2 ☐ Up to half of students 3 ☐ More than half of students
64a. Which of the following best describes the culture of the classroom:	1 ☐ Most students are resistant to learning. 2 ☐ Most students are compliant. 3 ☐ Most students are keenly motivated to learn.

C. ENGAGEMENT AND MOTIVATION

How characteristic are the following descriptions of the classroom?

65. Was the teacher passionate about what he or she was doing/teaching? *Passionate: enthusiastic, excited, invested, energetic, genuine, "on fire"*	1 ☐ Yes 0 ☐ No
66. The majority of students were <u>on-task</u> for almost all of the class period.	1 ☐ Yes 0 ☐ No
66a. How many students struggled to stay on task during the class period?	1 ☐ None to a few 2 ☐ About half 3 ☐ Most to all
67. How many students appear to be <u>actively engaged</u> in the lesson (beyond listening or reading, student were: raising hands, participating in discussion, writing, or drawing in relation to the lesson)?	1 ☐ None to a few 2 ☐ About half 3 ☐ Most to all
68. Do the majority of students appear to be <u>keenly interested</u> in the lesson or taking ownership of their learning?	1 ☐ Yes 0 ☐ No
69. Did the students play <u>leadership</u> roles, either voluntarily or from being encouraged by the teacher? **MARK ALL THAT APPLY**	1 ☐ Students volunteer for tasks 2 ☐ Students help peers without being asked 3 ☐ Students lead/facilitate discussions or discussion protocols; act as moderator 4 ☐ Students manage materials 5 ☐ None of the above
70. Which of the following behaviors or character traits did the students demonstrate/exhibit? **MARK ALL THAT APPLY**	1 ☐ Perseverance (*continuing to focus even when work is hard)* 2 ☐ Risk Taking/Academic Courage (*willing to self-evaluate, willing to disagree with peers)* 3 ☐ Collaboration *(giving peers feedback or help; use of respectful language)* 4 ☐ Craftsmanship *(demonstrating high level of skill)* 5 ☐ Curiosity *(asking questions, above and beyond the usual)* 6 ☐ Resistance *(refusing to cooperate, engage, or participate)* 7 ☐ Apathy *(not causing problems but not paying attention, engaging, or participating)* 8 ☐ Frustration *(trying to participate but getting impatient)* 9 ☐ Giving up *(trying at first, but giving up when things get hard)*

	10 ☐ None of the above

SECTION 3 NOTES:

This page has been left blank for double-sided copying.

Appendix E
Supplemental information for teacher instructional practice analyses

This page has been left blank for double-sided copying.

This appendix presents supplemental information for the analysis of teacher instructional practices. We first present information on the methodology used to develop the analytic measures of teacher instructional practices. We then present additional information on the methodology used to evaluate the intervention.

A. Supplemental information on the methodology used to develop teacher instructional practice measures

We developed empirical measures of teacher instructional practices through exploratory and confirmatory factor analyses using the teacher survey and observation data. We first describe the steps in this process: conceptual framework to group items into constructs, exploratory factor analysis (EFA) to determine optimal number of factors for each construct, confirmatory factor analysis (CFA) to test the fit of the model, and computation of construct scores for each teacher. We then discuss the predictive validity and internal consistency of these constructs.

1. Development of teacher instructional practice measures

We used a statistical method called factor analysis to develop the teacher instructional practice measures; this is a standard approach to develop measures from a set of individual items and establish their reliability and validity. We used both EFA and CFA, which are based on the common factor model and are typically conducted in conjunction with one another. Exploratory factor analysis is used to explore and refine the factor structure of the measures, whereas confirmatory factor analysis is used at the later stage of instrument development (after the testing instrument has been analyzed via exploratory factor analysis) for testing and validation purposes.

a. Conceptual framework

As the first step, we created a conceptual framework for the survey and observation items and their relationship to the unobserved (latent) constructs of teacher instructional practices these items were designed to measure. We estimated a total of 31 constructs (factors) measuring different aspects (topics) of 16 teacher instructional practices. Each construct was derived from one data source, either survey or observation. Some of the teacher instructional practices were measured by more than one construct (factor); see Table E.1.

Table E.1. Teacher instructional practice constructs by topic

Topic	Construct	Data source
Academic vocabulary	Academic vocabulary	Observation
Classroom climate	Classroom climate	Observation
Classroom management	Observed classroom management	Observation
	Self-reported classroom management	Survey
	Classroom disruption	Survey
Connected lessons	Stated learning purpose	Observation
	Connected lessons	Observation
Connections to world	Learning connected to personal experiences or real world	Observation
Content knowledge	Content knowledge development	Observation
	Demonstrate content knowledge in writing	Survey
Higher-order thinking	Observed higher-order thinking	Observation
	Higher-order thinking and reading	Survey
	Assessment of higher-order thinking	Survey
Multimedia use	Use of multimedia in instruction	Observation
	Online writing	Survey
Multiple types and purposes of writing	Engagement in multiple types of writing	Observation
	Writing for multiple purposes and audiences	Survey
Prior knowledge	Student prior knowledge	Survey
Student responsibility for learning	Student responsibility for their learning	Survey
Student independence	Student independence	Survey
Student participation	Collaborative discussion practices	Observation
	Student participation in discussion	Observation
	Student engagement in class	Observation
Use of text evidence	Text evidence use and argument writing	Observation
	Assessment of text evidence use	Survey
	Self-reported text evidence use	Survey
Reading, writing, and/or speaking about texts	Close reading and writing that demonstrates understanding of texts	Observation
	Reading, writing, and/or speaking about texts in past two weeks	Survey
	Reading, writing, and/or speaking about texts in typical week	Survey
Writing conventions	Use of writing conventions	Survey
	Feedback on writing conventions	Survey

b. Exploratory factor analysis

Exploratory factor analysis is a statistical analysis method typically used for the investigation of construct validity in cases where the relationships among variables are unknown or ambiguous (Atkinson et al. 2011). Because the items mapping to the same construct could measure several distinct dimensions of the same construct, we employed EFA to identify the underlying latent structure and dimensionality of the set of items (that is, the optimal number of factors these items are measuring). We employed EFA with Oblimin rotation for categorical outcomes to determine the latent dimensionality of the constructs. Oblimin rotation was used because all of the teacher instructional practices were assumed to be related and this rotation method allows the factors to correlate. To adjust for standard errors due to clustering of observations and survey responses for

the same teacher within a school year, we used a standard Huber-White estimation procedure that takes into account clustering of responses and assumes independence only among cluster units (teachers) but not the individual units (responses from the same teacher). We determined dimensionality structure of each construct by triangulating four pieces of information:

1. Statistical fit of the models with a varying number of factors to the data (based on the Bayesian predictive posterior *p*-value that should be at least 0.05 for a good fitting model; see Section C for more details).

2. Clean factor structure: factor loadings above 0.32, which equate to approximately 10 percent overlapping variance with the other items in that factor (Tabachnik and Fidel 2001); no or few item crossloadings on other factors; and no factors with a single item.

3. Substantive meaning and interpretability of the factors.

4. Adequate internal consistency reliability of the factors as a whole: ≥ 0.50, according to What Works Clearinghouse (WWC) standards.

Throughout the entire model-building and model-selection processes, we ensured that the models suggested by the EFA were consistent with our expectations of the conceptual relationship between the items and factors.

c. Confirmatory factor analysis

After we established the optimal number of factors for each construct, we fit a categorical CFA model to test whether the data supported the implied model (as driven by the EFA or/and theory).

To properly account for the multilevel structure of our data (observations and survey responses for each teacher in the fall and spring of the same school year are "nested" within that teacher), we opted to fit a multilevel categorical CFA model. Most generally, CFA models statistical relationships between observed variables and latent variables. In the most simplified form, for observation or survey response *i* for teacher *j,* this model can be specified as follows:

(V1) $$Y_{ij} = \beta_{oj} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \beta_{3j}X_{3ij} + \varepsilon_{ij},$$

where β_{oj} is the intercept coefficient (representing the mean level of estimated teacher instructional practice), β_{1j}, β_{2j}, and $\beta_{3j}$are the slope coefficients (or factor loadings) describing the effect of a unit change on items X_1, X_2, and X_3 in fall or spring observation or survey instrument *i* for teacher *j*, and ε_{ij} represents error in predicting Y_{ij} from known values of X_{ij}. We held the slopes (or factor loadings) constant across teachers. The intercepts (means) of each item were allowed to vary across clusters (teachers) and measure the distance from each teacher's intercept to the overall intercept β_o. Note that the notation does not account for the categorical

nature of the indicators, which have multiple intercepts (thresholds) that capture transition between response categories for binary-categorical items.[24]

We used an iterative approach to estimate factor models, where we first tested the hypothesized model structure as suggested by the EFA along with our theory of the relationship between the items and the factors and then continued to make necessary revisions to the model to obtain the optimal fit to the data and maximize internal consistency of the resulting factors (for example, by excluding items with nonsalient factor loadings < 0.32). To account for the small sample size, we used the Bayesian estimation method, which tends to outperform the standard Maximum Likelihood estimator in small samples (see, for example, Park and Yu 2018).

M*plus*—the statistical package we use for factor analysis—uses vague informative priors ~N(0, 10^{10}) for location parameters (that is, factor loadings and intercepts/thresholds) and noninformative priors for variances $p(\sigma^2) \propto 1$ as defaults. As mentioned previously, using these priors in small samples facilitates model identification (see van Erp et al. 2018). For each model we also evaluated model fit and convergence criteria for each parameter to determine whether we need to change priors' specifications.

For all 31 constructs, an underlying CFA models provided excellent fit to the data in the one-year sample according to the Bayesian PPP statistic (Table E.2)[25,26.] We also inspected Bayesian posterior distribution graphs to ensure convergence of model parameters.

[24] For more details about the multilevel CFA models, including the decomposition of the variance–covariance matrix within and between teachers, see Hox (2010).

[25] Results from the EFA analysis are available upon request.

[26] The PPP statistic controls the Type I error rate and detects structural misspecifications of the model and a 95 percent confidence interval for the difference in the f-statistic between the observed and replicated data. PPP values ≥ 0.05 typically indicate a good fit of the model to the data and values ≥ 0.50 indicate an excellent fit to the data. An excellent-fitting model is a PPP value around 0.5 and an f-statistic should include the value of zero falling close to the middle of the confidence interval (Muthén and Asparouhov 2012).

Table E.2. Statistical fit of the confirmatory factor analysis models for teacher instructional practices (one-year sample)

Teacher instructional practices	PPP	95% C.I.	
Academic vocabulary	0.50	–12.51	11.37
Classroom management	0.17	–25.64	81.06
Classroom climate	0.45	–28.19	27.91
Connected lessons	0.50	–22.40	20.92
Connections to world	0.46	–23.20	27.70
Content knowledge	0.41	–23.82	29.48
Higher-order thinking	0.25	–52.87	106.82
Multimedia use	0.50	–24.80	25.89
Multiple writing	0.48	–20.32	21.22
Student prior knowledge	0.48	–12.17	13.95
Student independence	0.49	–20.78	22.91
Student participation	0.44	–39.73	43.27
Student responsibility for learning	0.53	–17.56	15.78
Use of text evidence	0.47	–35.90	40.26
Writing about texts	0.36	–58.95	77.93
Writing conventions	0.19	–23.07	59.43

Source: Mathematica analysis of the one-year EL Education sample data using categorical confirmatory factor analysis with Bayesian estimation method.

PPP = predictive posterior *p*-value; PPP values > 0.5 indicate good fit of the model to the data.

95% C.I. = confidence interval for the difference in the f-statistic between the observed and replicated data.

In the one-year sample—which we used to develop the constructs—all constructs had high construct validity and internal consistency reliability. Construct validity, measured by the factor loadings, means that measures were strongly related to the underlying latent construct they were hypothesized to measure. Internal consistency reliability means that the items hypothesized to measure the same construct produce similar scores. We measured internal consistency reliability using McDonald's omega. This approach relies on the estimates from a CFA model that decomposes observed score variance into the true score and error variances and computes reliability as the ratio of true variance of the scores (signal) and the total variance of the scores (signal + noise):

(V2) $$\boldsymbol{\omega} = \frac{\sigma^2_{true\ variance}}{\sigma^2_{true\ variance} + \sigma^2_{error\ variance}}.$$

Omega reliability accounts for the differential strength of the relationship between each item and a factor and therefore provides a more accurate estimate of reliability than coefficient alpha (Zinbarg et al. 2005; Gadermann et al. 2012). There is no universal cutoff for a minimally acceptable value of internal consistency reliability statistic (such as alpha or omega), but WWC considers an internal consistency of 0.5 or higher as meeting the standards. In applied research, an internal consistency of 0.70 and above is typically considered desirable. Table E.3 provides additional details about internal consistency reliability of each individual construct.

Table E.3. Internal consistency by teacher practice construct

Teacher instructional practice construct	Teacher instructional practice sub-construct	Number of items	Internal consistency (omega)
Academic vocabulary	—	2	0.76
Classroom management	Observed classroom management	3	0.98
	Self-reported classroom management	6	0.84
	Classroom disruption	3	0.79
Classroom climate	—	7	0.90
Connected lessons	Stated learning purpose	2	0.78
	Connected lessons	2	0.62
Connections to world	—	3	0.83
Content knowledge	Content knowledge development	3	0.98
	Demonstrate content knowledge in writing	2	0.88
Higher-order thinking	Observed higher-order thinking	7	0.84
	Higher-order thinking and reading	7	0.89
	Assessment of higher-order thinking	4	0.89
Multimedia use	Use of multimedia in instruction	3	0.54
	Online writing	2	0.92
Multiple texts	—	2	0.24*
Multiple writing	Engagement in multiple types of writing	2	0.73
	Writing for multiple purposes and audiences	2	0.52
Prior knowledge	—	2	0.55
Student independence	—	4	0.73
Student participation	Collaborative discussion practices	4	0.72
	Student participation in discussion	2	0.75
	Student engagement in class	3	0.87
Student responsibility for learning	—	3	0.74
Use of text evidence	Text evidence use	4	0.87
	Assessment of text evidence use	2	0.87
	Self-reported text evidence use	2	0.79
Writing about texts	Reading, writing, and/or speaking about texts (observed)	5	0.81
	Reading, writing, and/or speaking about texts in past two weeks	4	0.81
	Reading, writing, and/or speaking about texts in a typical week	5	0.84
Writing conventions	Use of writing conventions	2	0.90
	Feedback on writing conventions	6	0.92

Source: Mathematica analysis of the one-year EL Education sample data using categorical confirmatory factor analysis with Bayesian estimation method.

*Does not meet the What Works Clearinghouse standard for sufficient internal consistency reliability (≥ 0.5).

d. Calculation of factor scores

For each teacher record (per wave and school year), we computed scores on each of the 31 constructs. For these summary scores to be reliable and valid representations of how the teacher

implements instructional practices, they should capture differential reliability of the items measuring these practices so that items with higher reliability and validity would get more weight in a composite score (see Table E.4). These reliability-weighted composite scores derived using the CFA procedure are referred to as "factor scores." Conceptually, a factor score is an estimate of the score that would have been observed for a teacher if it had been possible to measure the underlying construct (teacher instructional practice) directly. Factor scores represent the sum of the factor loadings of all observed items in a factor. These weights are estimated from the correlations between the items and reflect the degree of validity and reliability with which each survey or observation item measures a corresponding teacher instructional practice.

Reliability-weighted factor scores were computed by multiplying the observed score for each survey or observation item by the item's factor loading and summing the products of all items within a factor. We conducted two transformations to make the comparisons of teachers' scores across multiple constructs directly comparable and independent from the number of items in a factor and the scale of individual items (which could have different scales both within and across factors). First, we normalized each item to be on a scale from 0 to 1 to ensure the scores from different items always have the same metric. This is obtained by subtracting a minimum score from each teacher's observed score and dividing it by the difference between the maximum and minimum score of each item:

(V3) $$X_{normalized} = \frac{X - X_{minimum}}{X_{maximum} - X_{minimum}}$$

Second, we normalized the factor loadings within each factor and set their sum to 1 by dividing each factor loading by the sum of all loadings in a factor. Resulting scores always range between 0 and 1 regardless of the practical range for each survey or observation item and the number of items within a construct. As a result, resulting scores will always be directly comparable for teachers across different constructs. We required that a teacher had nonmissing responses to every survey or observation items within a construct to compute their score on that construct.

2. Predictive validity of teacher instructional practice measures

When examining outcomes of teachers' instructional practices, it is important to determine whether the instructional practices are associated with changes in student achievement. To determine the strength of this relationship we conducted two sets of analyses: (1) correlational analyses between each outcome and average standardized ELA scores at post-test, and (2) regression analyses of construct composite scores and test scores.

Table E.4. Item-level factor loadings by construct

Topic	Construct	Item	One-year sample loading	Two-year sample loading
General instruction				
Connected lessons	Connected lessons	Extent to which the overall lesson had activities that were connected to one another.	0.79	0.93
		Whether the lesson included a summary at the end to synthesize learning.	0.55	0.38
	Stated learning purpose	Whether the teacher referred back to the purpose/learning goals (that is, restating or pointing to purpose on the board) during the lesson.	0.80	0.91
		Teacher and/or students explicitly stated the purpose of instruction out loud (that is, Learning Target or Goal, "I Can" statement, Students Will Be Able To…).	0.80	0.94
Connections to world	Learning connected to personal experiences or real world	Students engaged in reading activities that involved connecting text to life during the class period with or without the teachers' help.	0.81	0.41
		Teacher provided instruction on academic vocabulary by using a real world or personal example to define a word (making connections to background knowledge).	0.77	0.54
		Whether the teacher made connections to the real world (that is, the usefulness or relevance of the lesson or topic to real world contexts beyond the classroom).	0.80	0.61
Content knowledge	Content knowledge development	Extent to which teachers engaged students in building their content knowledge.	0.97	0.97
		The approximate proportion of the class period the teacher spent on developing students' content knowledge.	0.97	0.97
		The degree to which the content knowledge instruction was connected to the main point of the lesson.	0.98	0.98
	Demonstrate content knowledge in writing	The extent to which the demonstration of content knowledge or understanding of text is important when grading or evaluating students' writing.	0.87	0.85
		The extent to which the demonstration of content knowledge or understanding of text is important when providing students with feedback on their writing.	0.89	0.91
Higher-order thinking	Assessment of higher-order thinking	The extent to which the development of ideas and use of evidence to support ideas is important when grading or evaluating students' writing.	0.81	0.86
		The extent to which the quality of writing important when grading or evaluating students' writing	0.85	0.94

Topic	Construct	Item	One-year sample loading	Two-year sample loading
		The extent to which the development of ideas and use of evidence to support ideas is important when providing students with feedback on their writing.	0.72	0.82
		The extent to which the quality of thinking is important when providing students with feedback on their writing.	0.89	0.95
	Higher-order thinking and reading	The frequency with which teachers focus on analysis of individuals, events, and ideas in the text (for example, how they are introduced, how they interact with each other, and how text makes connections/distinctions among them) when students read texts in class.	0.60	0.81
		The frequency with which teachers focus on analyzing differences and similarities in perspectives between multiple texts on the same topic when students read texts in class.	0.85	0.86
		The frequency with which teachers focus on critiquing or evaluating arguments and specific claims in context when students read texts in class.	0.60	0.86
		The frequency with which teachers focus on determining central idea(s) or theme(s) of the text when students read texts in class.	0.74	0.88
		The frequency with which teachers focus on determining the author's point of view or purpose when students read texts in class.	0.74	0.95
		The frequency with which teachers focus on integrating or comparing and contrasting information in different media or formats when students read texts in class.	0.86	0.86
		The frequency with which teachers focus on integrating or interpreting information across sections of the text when students read texts in class.	0.74	0.78
	Observed higher-order thinking	Extent to which the teacher asked higher-order thinking questions either verbally or in writing.	0.93	0.96
		Extent to which the teacher reinforced higher-order thinking in class.	0.90	0.96
		Students engaged in reading activities that involved making predictions during the class period with or without the teachers' help.	0.34	0.40
		Teacher asked students to make inferences, analyze, synthesize, draw conclusions, evaluate, and/or critique either verbally or in writing.	0.74	0.90
		Teacher asked students to make predictions either verbally or in writing.	0.51	0.56
		Teacher asked students questions to encourage them to think at a higher level (see definition).	0.65	0.59

Topic	Construct	Item	One-year sample loading	Two-year sample loading
Multimedia use	Online writing	Teacher asked the student probing questions, for more information, or follow-up questions.	0.45	0.73
		Frequency with which teacher asks students to create videos or webcasts.	0.92	0.95
		Frequency with which teacher asks students to write for social networking, blogs, or wikis.	0.92	0.95
	Use of multimedia in instruction	Teacher used chalk board, smart board, projector, or chart paper as an instructional format in the class period.	0.53	0.69
		Teacher used games, role play, and drama as an instructional format in the class period.	0.50	0.41
		Teacher used multimedia, such as films, videos, music, and art, as an instructional format in the class period.	0.57	0.69
Prior knowledge	Student prior knowledge	In the past two weeks, frequency with which teachers asked students to relate what they read to their own experience or to something they have learned before.	0.46	0.13
		When students read texts in class, how often teachers focused on relating the story or literary work, its characters, and/or its themes to something they have read before.	0.76	0.42
Student independence	Student independence	The frequency with which students asked for more challenging work in the past two weeks.	0.65	0.64
		The frequency with which students checked their own progress against learning targets in the past two weeks.	0.66	0.59
		The frequency with which students gave input in setting learning targets in the past two weeks.	0.63	0.55
		The frequency with which students participated in developing rubrics or evaluation criteria against learning targets in the past two weeks.	0.60	0.75
Student participation	Collaborative discussion practices	Frequency with which students responded to peer comments and/or build on each other's thinking.	0.87	0.92
		Teacher used student pair or group work as an instructional technique in the class period.	0.46	0.58
		Teacher asked other students to help answer the question in discussions.	0.34	0.76
		The thoughtfulness of the classroom discussion based on students' responses: (1) students' responses were automatic or rote; (2) students sometimes responded thoughtfully; and (3) students sometimes responded thoughtfully, plus generated their own probing questions.	0.78	0.96

Topic	Construct	Item	One-year sample loading	Two-year sample loading
Student responsibility for learning	Student engagement in class	The approximate proportion of students who struggled to stay on task during the class period.	0.90	0.85
		Whether the majority of students were on task for almost all the class period.	0.91	0.94
	Student participation in discussion	The extent to which students participated in class.	0.76	0.95
		The extent to which students spoke in the classroom.	0.79	0.89
	Student responsibility for their learning	The frequency with which students provided feedback on each other's work in the past two weeks.	0.53	0.00
		The frequency with which students shared their work with their peers in the past two weeks.	0.85	0.21
		The frequency with which students worked with other students in the past two weeks.	0.70	0.66
Reading and writing instruction				
Academic vocabulary	Academic vocabulary	Degree to which the teacher focused on students' academic vocabulary instruction.	0.76	0.79
		Students engaged in reading activities that involved determining word meanings from text(s) during the class period with or without the teachers' help.	0.81	0.83
Multiple types and purposes of writing	Engagement in multiple types of writing	Count of writing purposes, including: practicing grammar, mechanics, and/or spelling; literary; marking arguments based on evidence; writing to demonstrate the following of text; promote personal reflection on the text; synthesize what students know; learn content knowledge; record information or ideas; write from sources; literary analysis; analyze text structure; and others.	0.87	0.97
		Number of writing activities in different genres or modes of the writing that the students engaged in. Genres or modes can include poem, short story, biography/memoir/personal essay, narrative, descriptive, informational/expository, opinion writing, literary analysis, book review or literature review, persuasion (no sources or references used), argument writing (sources or references used), free writing, graphic organizer, letters, speeches, and other genre or modes of writing.	0.63	0.95
	Writing for multiple purposes and audiences	Number of the following types of audiences teachers asked students to write for: family members; school officials; teachers; other students; community members; and community organizations, government officials, or businesses.	0.51	0.37

Topic	Construct	Item	One-year sample loading	Two-year sample loading
Reading, writing, and/or speaking about texts		Number of the following types of writing teachers asked students to do in the past two weeks: writing to inform, writing to argue or present an opinion, and writing to convey experience in narrative form.	0.68	0.39
	Close reading and writing that demonstrates understanding of text	Purpose of the writing is to demonstrate understanding of the text(s).	0.50	0.80
		Purpose of the writing was to cite sources or evidence.	0.82	0.54
		Students engaged in close reading (students reread text to look for answers to specific questions) during the class period with or without the teachers' help.	0.54	0.88
		Students reviewed text to find evidence to support a response or viewpoint.	0.91	0.82
		The extent to which the reading activities focused on the meaning of the text (for example, main idea, key information, main theme, characters, and plot).	0.55	0.87
	Reading, writing, and/or speaking about texts in past two weeks	Frequency with which students discussed texts they read with partners or a small group of students in the past two weeks.	0.78	0.80
		Frequency with which students produced extended writing by incorporating key details from texts they read in the past two weeks.	0.57	0.68
		Frequency with which students shared their ideas about and/or understanding of the texts they read with the whole class in the past two weeks.	0.76	0.86
		Frequency with which students wrote about texts they read in the past two weeks.	0.76	0.91
	Reading, writing, and/or speaking about texts in typical week	Frequency with which teachers ask students to do a group activity or project about what they have read in a given week.	0.62	0.56
		Frequency with which teachers ask students to explain or support their understanding of what they have read in a given week.	0.73	0.69
		Frequency with which teachers ask students to write about something they have read in a given week.	0.80	0.88
		Frequency with which teachers give students time to read books they have chosen themselves in a given week.	0.87	0.89
		Frequency with which teachers ask students to describe the style or structure of the text they have read in a given week.	0.56	0.48
Use of text evidence	Assessment of text evidence use	The extent to which the accuracy or appropriateness of students' references to texts important for text-based writing when grading or evaluating students' writing.	0.88	0.83

Topic	Construct	Item	One-year sample loading	Two-year sample loading
Writing conventions		The extent to which the accuracy or appropriateness of their references to texts important for text-based writing when providing students with feedback on their writing.	0.87	0.84
	Self-reported text evidence use	In the past two weeks, frequency with which teachers ask students to write citing evidence or information from a text that students read.	0.87	0.78
		In the past two weeks, frequency with which teachers ask students to write thinking tasks, such as graphic organizers, that capture students' thoughts on the text and relevant evidence.	0.75	0.91
	Text evidence use and argument writing	Purpose of the writing is making arguments based on evidence.	0.92	0.94
		Students engaged in argument writing (using sources and references).	0.93	0.81
		Students engaged in writing activities involving citing sources or evidence with or without the teacher's help.	0.72	0.94
		Teacher asked students to cite evidence from the text to support their responses either verbally or in writing.	0.58	0.72
	Feedback on writing conventions	The extent to which mechanics, conventions, paragraphing, and using appropriate formats are important when grading or evaluating students' writing.	0.87	0.80
		The extent to which the effective use of language (for example, sentence variety, word choice, and tone) is important when grading or evaluating students' writing.	0.79	0.91
		The extent to which the organization of ideas important when grading or evaluating students' writing.	0.79	0.98
		The extent to which mechanics and using appropriate formats important when providing students with feedback on their writing.	0.80	0.81
		The extent to which the effective use of language (for example, sentence variety, tone, and word choice) is important when providing students with feedback on their writing.	0.83	0.92
		The extent to which the organization of ideas is important when providing students with feedback on their writing.	0.81	0.98
	Use of writing conventions	Purpose of the writing was to practice grammar, mechanics, and/or spelling.	0.91	0.89
		Students engaged in writing activities that involved writing conventions (grammar, mechanics, spelling, paragraphing, and formatting—such as use of subheadings or quotations).	0.91	0.86

Topic	Construct	Item	One-year sample loading	Two-year sample loading
Classroom management and environment				
Classroom climate	Classroom climate	Approximate proportion of students who were receptive to the teacher's instructions and feedback.	0.91	0.84
		Approximate proportion of students who were resistant to learning, compliant, and keenly motivated.	0.86	0.60
		Degree to which teachers and students demonstrated respect for one another.	0.84	0.84
		Degree to which teachers successfully encouraged students to interact with one another.	0.40	0.92
		Degree to which teachers were warm, empathetic, and/or supportive.	0.70	0.88
		Whether the teachers were passionate about what they were doing or teaching.	0.83	0.90
Classroom management	Classroom disruption (reverse)	The extent to which student(s) disrupts class for 5 to 10 minutes (reverse coded).	0.90	0.95
		The extent to which student(s) disrupts class for more than 10 minutes (reverse coded).	0.87	0.95
		The extent to which transitions from one activity to another take more than 5 minutes.	0.42	0.79
	Observed classroom management	Degree to which teachers' attempts to redirect student disruptions were effective.	0.98	0.97
		Extent to which teachers enforced behavioral rules in the classroom.	0.97	0.95
		Whether teachers' redirection of student disruption took away from learning time.	0.97	0.98
	Self-reported classroom management	The extent to which teachers emphasize positive character traits and values in the classroom.	0.69	0.80
		The extent to which teachers manage their class well.	0.48	0.79
		The extent to which teachers often acknowledge positive student behavior.	0.70	0.87
		The extent to which teachers redirect students back to the topic when they get off-task.	0.72	0.76
		The extent to which teachers require that students in their classroom follow the rules at all times.	0.74	0.39
		The extent to which teachers set clear expectations for student behavior.	0.77	0.71

For the correlation analyses, we computed bivariate Pearson correlations between the standardized ELA post-test scores, aggregated at the teacher level, and each of the 31 constructs described previously, aggregated at the school year level (composed of the fall and spring scores) for teachers who had scores in both fall and spring. The correlation coefficient between the two sets of scores is also known as the validity coefficient.[27] The validity coefficients presented in Table E.5 indicate positive relationships between instructional practice outcomes and student ELA achievement, though relationships tended to be weak with the majority of correlation coefficients under 0.29 (Cohen 1988). However, all correlations above 0.13 were statistically significant.

Our second approach to estimate predictive validity involved ordinary least square regression, where we regressed the composite scores of the constructs onto the standardized ELA post-test scores. We repeated this test for each analytic sample, controlling for treatment status. Results of these regression analyses are shown in Table E.6. The R-squared statistics show a total percent of variation in the standardized ELA scores explained by the composites of the outcome measures. F-statistics provide a global test of significance for the predictor in the model. The results show a strong association between the observation constructs and the ELA post-test scores among teachers in the one-year sample, as well as associations between the survey constructs and the ELA post-test scores among teachers in the two-year sample.

[27] The validity coefficient (Pearson's r) can range from –1 to + 1. Large coefficients close to 1 in absolute value indicate high predictive validity of the new measure, coefficients close to 0 indicate weak predictive validity, and coefficients of 0 indicate no association between the scores.

Table E.5. Correlations between teacher instructional practice constructs and standardized English-language arts outcome scores

Topic	Construct	Observation	Survey	Correlation Coefficient	Sample size
General instruction					
Academic vocabulary	Academic vocabulary	x		0.07	241
Connected lessons	Connected lessons	x		0.23***	241
	Stated learning purpose	x		0.13**	241
Connections to world	Learning connected to personal experiences or real world	x		-0.10	241
Content knowledge	Content knowledge development	x		-0.03	241
	Demonstrate content knowledge in writing		x	0.04	280
Higher-order thinking	Assessment of higher order thinking		x	0.03	280
	Higher order thinking and reading		x	0.03	280
	Observed higher order thinking	x		0.18***	241
Multimedia use	Online writing		x	-0.04	280
	Use of multimedia in instruction	x		0.07	241
Prior knowledge	Student prior knowledge		x	0.00	280
Student independence	Student independence		x	-0.09	280
Student participation	Collaborative discussion practices	x		0.30***	241
	Student engagement in class	x		0.24***	241
	Student participation in discussion	x		0.25***	241
Student responsibility for learning	Student responsibility for their learning		x	0.14**	280
Reading and writing instruction					
Multiple types and purposes of writing	Engagement in multiple types of writing	x		0.04	241
	Writing for multiple purposes and audiences		x	-0.03	280
Reading, writing, and/or speaking about texts	Close reading and writing that demonstrates understanding of texts	x		0.10	241
	Reading, writing, and/or speaking about texts in past two weeks		x	0.17***	280
	Reading, writing, and/or speaking about texts in typical week		x	0.03	280

Topic	Construct	Observation	Survey	Correlation Coefficient	Sample size
Use of text evidence	Assessment of text evidence use		x	0.05	280
	Self-reported text evidence use		x	0.05	280
	Text evidence use and argument writing	x		0.02	241
Writing conventions	Feedback on writing conventions		x	-0.02	280
	Use of writing conventions		x	-0.03	241
Classroom management and environment					
Classroom climate	Classroom climate	x		0.37***	241
Classroom management	Classroom disruption		x	0.27***	280
	Observed classroom management	x		-0.04	241
	Self-reported classroom management		x	0.08	281

Source: Sample includes all teachers in the one-year and/or two-year study.

[a] The items in the classroom disruption construct were reverse coded to have a similar scale as other constructs (that is, higher scores are more positive instructional practices than lower scores). See Chapter V for more information about this construct.

***Indicates the correlation is significant at the 0.01 level.

**Indicates the correlation is significant at the 0.05 level.

Table E.6. Regression results of the validity of the teacher instructional practice composite outcomes

	Outcome	Sample size	R2	F statistic
One-year sample	Observation Composite	195	0.05	10.93***
One-year sample	Survey Composite	132	0.02	2.40
Two-year sample	Full Composite	39	0.03	1.17
Two-year sample	Observation Composite	46	0.04	1.67
Two-year sample	Survey Composite	39	0.13	5.62**

Source: For one-year observation outcomes, estimates used data from 2015–2016 for Cohort 2 schools and 2016–2017 for Cohort 3 schools. For one-year survey outcomes, estimates used data from spring 2016 for Cohort 2 schools and spring 2017 for Cohort 3 schools. For two-year outcomes, estimates used data from 2017–2018. The ELA post-test score data are from administrative educator records for 2016–2017.

Note: Effect size is the adjusted difference between treatment and control group average teacher instructional practice scores divided by the unadjusted pooled standard deviation.

***Indicates there are significant differences among the groups at $p < 0.01$.

**Indicates there are significant differences among the groups at $p < 0.05$.

3. Internal consistency of teacher instructional practice constructs

To further test the strength of our constructs, we measured internal consistency reliability—the degree to which the constructs related to one another and measure a similar concept. Here we estimated joint Cronbach's alpha statistics for the observation constructs, survey constructs, and combined survey and observation constructs (see Table E.7). Internal consistency of 0.50 or higher is considered sufficient by WWC. In applied research, internal consistency of 0.70 and above is typically considered desirable.

Table E.7. Internal consistency of teacher instructional practice measures

Analysis	Measures	Cronbach's alpha
One-year sample	Observation constructs	0.74
One-year sample	Survey constructs	0.78
Two-year sample	All constructs	0.82
Two-year sample	Observation constructs	0.70
Two-year sample	Survey constructs	0.82

Source: For one-year observation outcomes, estimates used data from 2015–2016 for Cohort 2 schools and 2016–2017 for Cohort 3 schools. For one-year survey outcomes, estimates used data from spring 2016 for Cohort 2 schools and spring 2017 for Cohort 3 schools. For two-year outcomes, estimates used data from 2017–2018. The ELA post-test score data are from administrative educator records for 2016–2017.

B. Additional information on the methodology used to evaluate the intervention

This section provides supplemental information and technical details about the analyses and findings discussed in Chapter V. We present details on the handling of missing data and baseline equivalence computation and data. We then present additional information on the impact estimation models.

1. Schools included in the analyses and cluster-level attrition

We conducted impact analyses using individual teachers as the unit of analysis, but we conducted random assignment at the cluster level by assigning schools to treatment or control. Because schools were randomly assigned, it is important to assess the degree to which the schools that were randomized are represented in the analyses to know whether the random assignment structure is intact.

There were 79 schools randomly assigned to treatment and control (40 to treatment and 39 to control) and within these schools there were 66 novice teachers (32 treatment and 34 control). The novice teacher sample is the focus of our primary research questions, and novice teacher status[28] was a characteristic determined prior to randomization and the introduction of the Teacher Potential Project (TPP); the definition of novice teacher status was applied consistently for treatment and control schools, and the study did not allow teachers to join the study after random assignment.

The novice teachers included in the analyses using observation data were from 44 schools (20 treatment and 24 control), which represents 33 percent overall attrition and 8.1 percent differential attrition (see Table E.8). All teachers included in analyses using observation data were from 53 schools (26 treatment and 27 control), which represents 33 percent overall attrition and 4.2 percent differential attrition.

Novice teachers included in analyses using self-reported data from teacher surveys were from 38 schools (18 treatment and 20 control), which represents 42 percent overall attrition and 2.6 percent differential attrition. All teachers who were included in analyses using self-reported survey data were from 51 schools (27 treatment and 24 control), which represents 35 percent overall attrition and 6.0 percent differential attrition. Some of the survey outcomes had different attrition rates, as indicated in the notes in Table E.8.

[28] Novice teachers are defined as teachers who have zero to four years of full-time teaching experience the first year of their participation in the study. These years of experience do not include student teaching or substitute teaching.

Table E.8. School-level attrition for one-year impact analyses

	Analytic sample size of schools			Attrition rates	
Sample	Treatment	Control	Total	Overall	Differential
Observation data					
Novice	20	24	44	33%	8.1%
All teachers	26	27	53	33%	4.2%
Survey data					
Novice[1]	18	20	38	42%	2.6%
All teachers[2]	27	24	51	35%	6.0%

Source: Administrative educator records for 2014-2015 for Cohort 1 pilot schools and 2015-2016 for Cohort 2 schools.

Note: Numbers of schools reported. The count of overall Ns in this table reflects the original numbers of schools at the time of random assignment. For novice teachers, the reference group numbers which were used as the denominators in the attrition rate calculations were based on the 66 schools that were randomly assigned that had novice teachers (32 treatment and 34 control). For all teachers, the reference group numbers were 79 schools (40 treatment and 39 control).

[1] The outcome, self-reported overall ELA instructional practice, had 50 percent overall attrition and no differential attrition. The outcomes, student responsibility for their learning and student independence had 45 percent overall attrition and 2.8 percent differential attrition. The outcome, reading, writing, and/or speaking about texts in the past two weeks, had 44 percent overall attrition and 2.8% differential attrition. The outcome, use of writing conventions, had 33 percent overall attrition and 8.1 percent differential attrition.

[2] The outcome, self-reported overall ELA instructional practice, had 44 percent overall attrition and 1.4 percent differential attrition. The outcome, use of writing conventions, had 33 percent overall attrition and 4.2 percent differential attrition.

2. Missing data

We conducted complete case analyses, meaning any teacher with missing baseline data and the outcome measure of interest was omitted from that analysis. For the analyses on the composite score outcomes, we required that the teacher had a score for each of the constructs that went into the composite at baseline and follow-up.

3. Assessing baseline equivalence of school- and classroom-level characteristics

We next evaluated the extent to which our treatment and comparison groups of each analytic sample were similar at baseline on observed characteristics at both the school- and classroom-levels. This step provides insight into the composition of our study samples, specifically how they compare among empirically important characteristics. We conducted a series of t-tests, calculating effect sizes as the difference in means divided by the pooled standard deviation. School characteristics include charter status, number of full-time teachers, percent of students who were eligible for free or reduced-price meals, and total number of students from the Common Core of Data. Classroom-level characteristics include the percent of students who were black non-Hispanic, white non-Hispanic, other race, Hispanic, female, receiving special education services, English-language learners, receiving free or reduced-price lunch, as well as average scores on math and ELA pre-tests from student administrative data. For the one-year and two-year studies, we also measured differences in the baseline version of each outcome for teachers for whom we observed that outcome.

For each t-test we computed descriptive statistics by group (sample size, means, and standard deviations) and the corresponding t-statistic, *p*-value, and effect size (Hedges' g), along with the 95 percent confidence interval. Summary statistics for school-level characteristics are provided in Table V.1 in the main text. The tables include the results of the t-tests for the classroom characteristics for novice teachers and for the full sample of teachers (novice and experienced) in each analytic sample. For the one-year observation analysis samples, we measured for differences in characteristics for Cohort 2 and 3 teachers who had outcome data (each outcome is equal to the sum of the fall and spring scores for each measure). In addition to the characteristics described previously, for the one-year survey samples we measured for differences in the baseline version of the outcome (measured in the fall) for Cohort 2 and 3 teachers who had outcome data in the spring. For the two-year QED sample, we measured baseline differences from the 2016–2017 school year for teachers who had outcome data (data from the 2017-18 school year) using data from fall 2016 as the baseline measure.

On average, novice teachers in the treatment group of the one-year observation analytic samples had more special education students in their classrooms than novice teachers in the comparison group (Table E.9). Teachers did not differ significantly in any other dimension.

Comparison group teachers in the one-year analytic samples of survey outcomes had more white non-Hispanic students and fewer black non-Hispanic students in their classrooms than teachers in the treatment group; however, these difference were only statistically significant among the full sample of teachers for select outcomes (Table V.10). Teachers in the treatment group tended to have more special education students in their classrooms than teachers in the comparison group, and these differences were larger and more statistically significant among novice teachers. Teachers in the treatment group also had higher scores on assessment of higher-order thinking (novice only) and, (2) self-reported text evidence use (novice and full). Conversely, scores on use of writing conventions baseline were lower among treatment teachers in both the novice and full analytic samples and classroom disruption scores were lower among treatment teachers in the novice analytic sample.

In several of the two-year analytic samples, the average teacher instructional practice baseline score was significantly higher among teachers in the treatment group (Table E.11). Teachers did not differ significantly along other dimensions in the two-year sample.

Table E.9. Comparison of characteristics of teachers' students at baseline, one-year observation sample

Covariate	t-statistic	Effect size	Treatment			Comparison		
			Mean	SD	N	Mean	SD	N
Novice teachers								
Standardized ELA pre-test score	1.31	0.29	-0.64	0.48	37	-0.50	0.47	43
Standardized math pre-test score	1.02	0.23	-0.63	0.48	37	-0.52	0.49	43
Male	-0.05	0.01	0.53	0.08	37	0.53	0.07	43
White non-Hispanic	1.19	0.27	0.28	0.33	37	0.36	0.31	43
Black non-Hispanic	-1.16	0.26	0.48	0.38	37	0.39	0.30	43
Other race	0.26	0.06	0.06	0.06	37	0.06	0.06	43
Hispanic	-0.02	0.00	0.19	0.25	37	0.19	0.18	43
Free or reduced-price lunch eligibility	-0.56	0.13	0.76	0.25	37	0.73	0.24	43
English-language learner	-0.71	0.16	0.06	0.09	37	0.05	0.08	43
Special education status	-3.15***	0.67	0.14	0.08	37	0.08	0.07	43
All teachers								
Standardized ELA pre-test score	0.51	0.08	-0.52	0.51	75	-0.49	0.49	89
Standardized math pre-test score	1.11	0.17	-0.57	0.49	75	-0.48	0.54	89
Male	-0.8	0.13	0.52	0.11	75	0.51	0.07	89
White non-Hispanic	1.25	0.20	0.25	0.30	75	0.31	0.34	89
Black non-Hispanic	-1.64	0.26	0.50	0.36	75	0.42	0.32	89
Other race	-0.63	0.10	0.07	0.08	75	0.06	0.07	89
Hispanic	0.97	0.15	0.18	0.22	75	0.21	0.19	89
Free or reduced-price lunch eligibility	0.61	0.10	0.75	0.27	75	0.77	0.21	89
English-language learner	-0.28	0.04	0.07	0.11	75	0.07	0.14	89
Special education status	-1.06	0.17	0.12	0.11	75	0.11	0.10	89

Source: Administrative educator records for 2014-2015 for Cohort 1 pilot schools and 2015-2016 for Cohort 2 schools.

Baseline statistics are presented for the one-year analytic sample, which includes teachers who had all baseline and observation construct (outcome) data. The analytic sample did not differ across observation construct analyses because no outcome data were missing.

*Statistically significant at the 5 percent level, two-tailed test.

**Statistically significant at the 1 percent level, two-tailed test.

*** Statistically significant at the .01 percent level, two-tailed test.

ELA = English language arts; SD = standard deviation.

Table E.10. Comparison of characteristics of novice teachers at baseline, one-year observation sample

	Construct	Covariate	t-statistic	Effect size	Treatment			Comparison		
					Mean	SD	N	Mean	SD	N
General Instruction										
Connected lessons	Connected lessons	Connected lessons	-0.67	0.17	0.52	0.23	31	0.48	0.21	34
		Standardized ELA pre-test score	1.12	0.28	-0.57	0.46	31	-0.45	0.44	34
		Standardized math pre-test score	1.28	0.32	-0.57	0.46	31	-0.43	0.46	34
		Male	-0.16	0.04	0.52	0.08	31	0.52	0.06	34
		White non-Hispanic	1.16	0.29	0.29	0.33	31	0.39	0.33	34
		Black non-Hispanic	-1.00	0.25	0.46	0.37	31	0.38	0.32	34
		Other race	0.05	0.01	0.06	0.06	31	0.06	0.06	34
		Hispanic	-0.17	0.04	0.19	0.25	31	0.18	0.18	34
		Free or reduced-price lunch eligibility	0.05	0.01	0.73	0.26	31	0.74	0.25	34
		English-language learner	-0.45	0.11	0.06	0.07	31	0.05	0.06	34
		Special education status	-2.00**	0.49	0.13	0.09	31	0.10	0.07	34
	Stated learning purpose	Stated learning purpose	-3.48***	0.80	0.64	0.35	31	0.34	0.36	34
		Standardized ELA pre-test score	1.12	0.28	-0.57	0.46	31	-0.45	0.44	34
		Standardized math pre-test score	1.28	0.32	-0.57	0.46	31	-0.43	0.46	34
		Male	-0.16	0.04	0.52	0.08	31	0.52	0.06	34
		White non-Hispanic	1.16	0.29	0.29	0.33	31	0.39	0.33	34
		Black non-Hispanic	-1.00	0.25	0.46	0.37	31	0.38	0.32	34
		Other race	0.05	0.01	0.06	0.06	31	0.06	0.06	34
		Hispanic	-0.17	0.04	0.19	0.25	31	0.18	0.18	34
		Free or reduced-price lunch eligibility	0.05	0.01	0.73	0.26	31	0.74	0.25	34
		English-language learner	-0.45	0.11	0.06	0.07	31	0.05	0.06	34
		Special education status	-2.00**	0.49	0.13	0.09	31	0.10	0.07	34
Connections to world	Learning connected to personal experiences or real world	Learning connected to personal experiences or real world	-0.03	0.01	0.26	0.29	31	0.25	0.26	34
		Standardized ELA pre-test score	1.12	0.28	-0.57	0.46	31	-0.45	0.44	34
		Standardized math pre-test score	1.28	0.32	-0.57	0.46	31	-0.43	0.46	34
		Male	-0.16	0.04	0.52	0.08	31	0.52	0.06	34
		White non-Hispanic	1.16	0.29	0.29	0.33	31	0.39	0.33	34
		Black non-Hispanic	-1.00	0.25	0.46	0.37	31	0.38	0.32	34
		Other race	0.05	0.01	0.06	0.06	31	0.06	0.06	34
		Hispanic	-0.17	0.04	0.19	0.25	31	0.18	0.18	34
		Free or reduced-price lunch eligibility	0.05	0.01	0.73	0.26	31	0.74	0.25	34
		English-language learner	-0.45	0.11	0.06	0.07	31	0.05	0.06	34
		Special education status	-2.00**	0.49	0.13	0.09	31	0.10	0.07	34

	Construct	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
Content knowledge	Content knowledge development	Content knowledge development	0.35	0.09	0.20	0.35	31	0.23	0.37	34
		Standardized ELA pre-test score	1.12	0.28	-0.57	0.46	31	-0.45	0.44	34
		Standardized math pre-test score	1.28	0.32	-0.57	0.46	31	-0.43	0.46	34
		Male	-0.16	0.04	0.52	0.08	31	0.52	0.06	34
		White non-Hispanic	1.16	0.29	0.29	0.33	31	0.39	0.33	34
		Black non-Hispanic	-1.00	0.25	0.46	0.37	31	0.38	0.32	34
		Other race	0.05	0.01	0.06	0.06	31	0.06	0.06	34
		Hispanic	-0.17	0.04	0.19	0.25	31	0.18	0.18	34
		Free or reduced-price lunch eligibility	0.05	0.01	0.73	0.26	31	0.74	0.25	34
		English-language learner	-0.45	0.11	0.06	0.07	31	0.05	0.06	34
		Special education status	-2.00**	0.49	0.13	0.09	31	0.10	0.07	34
Higher-order thinking	Observed higher order thinking	Observed higher order thinking	-2.32**	0.56	0.56	0.21	31	0.44	0.22	34
		Standardized ELA pre-test score	1.12	0.28	-0.57	0.46	31	-0.45	0.44	34
		Standardized math pre-test score	1.28	0.32	-0.57	0.46	31	-0.43	0.46	34
		Male	-0.16	0.04	0.52	0.08	31	0.52	0.06	34
		White non-Hispanic	1.16	0.29	0.29	0.33	31	0.39	0.33	34
		Black non-Hispanic	-1.00	0.25	0.46	0.37	31	0.38	0.32	34
		Other race	0.05	0.01	0.06	0.06	31	0.06	0.06	34
		Hispanic	-0.17	0.04	0.19	0.25	31	0.18	0.18	34
		Free or reduced-price lunch eligibility	0.05	0.01	0.73	0.26	31	0.74	0.25	34
		English-language learner	-0.45	0.11	0.06	0.07	31	0.05	0.06	34
		Special education status	-2.00**	0.49	0.13	0.09	31	0.10	0.07	34
Multimedia use	Use of multimedia in instruction	Use of multimedia in instruction	0.20	0.05	0.38	0.24	31	0.39	0.22	34
		Standardized ELA pre-test score	1.12	0.28	-0.57	0.46	31	-0.45	0.44	34
		Standardized math pre-test score	1.28	0.32	-0.57	0.46	31	-0.43	0.46	34
		Male	-0.16	0.04	0.52	0.08	31	0.52	0.06	34
		White non-Hispanic	1.16	0.29	0.29	0.33	31	0.39	0.33	34
		Black non-Hispanic	-1.00	0.25	0.46	0.37	31	0.38	0.32	34
		Other race	0.05	0.01	0.06	0.06	31	0.06	0.06	34
		Hispanic	-0.17	0.04	0.19	0.25	31	0.18	0.18	34
		Free or reduced-price lunch eligibility	0.05	0.01	0.73	0.26	31	0.74	0.25	34
		English-language learner	-0.45	0.11	0.06	0.07	31	0.05	0.06	34
		Special education status	-2.00**	0.49	0.13	0.09	31	0.10	0.07	34
Student participation	Collaborative discussion practices	Collaborative discussion practices	-0.74	0.19	0.53	0.16	31	0.49	0.21	34
		Standardized ELA pre-test score	1.12	0.28	-0.57	0.46	31	-0.45	0.44	34
		Standardized math pre-test score	1.28	0.32	-0.57	0.46	31	-0.43	0.46	34
		Male	-0.16	0.04	0.52	0.08	31	0.52	0.06	34
		White non-Hispanic	1.16	0.29	0.29	0.33	31	0.39	0.33	34
		Black non-Hispanic	-1.00	0.25	0.46	0.37	31	0.38	0.32	34
		Other race	0.05	0.01	0.06	0.06	31	0.06	0.06	34

	Construct	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
		Hispanic	-0.17	0.04	0.19	0.25	31	0.18	0.18	34
		Free or reduced-price lunch eligibility	0.05	0.01	0.73	0.26	31	0.74	0.25	34
		English-language learner	-0.45	0.11	0.06	0.07	31	0.05	0.06	34
		Special education status	-2.00**	0.49	0.13	0.09	31	0.10	0.07	34
	Student engagement in class	Student engagement in class	2.69***	0.64	0.52	0.31	31	0.68	0.18	34
		Standardized ELA pre-test score	1.12	0.28	-0.57	0.46	31	-0.45	0.44	34
		Standardized math pre-test score	1.28	0.32	-0.57	0.46	31	-0.43	0.46	34
		Male	-0.16	0.04	0.52	0.08	31	0.52	0.06	34
		White non-Hispanic	1.16	0.29	0.29	0.33	31	0.39	0.33	34
		Black non-Hispanic	-1.00	0.25	0.46	0.37	31	0.38	0.32	34
		Other race	0.05	0.01	0.06	0.06	31	0.06	0.06	34
		Hispanic	-0.17	0.04	0.19	0.25	31	0.18	0.18	34
		Free or reduced-price lunch eligibility	0.05	0.01	0.73	0.26	31	0.74	0.25	34
		English-language learner	-0.45	0.11	0.06	0.07	31	0.05	0.06	34
		Special education status	-2.00**	0.49	0.13	0.09	31	0.10	0.07	34
	Student participation in discussion	Student participation in discussion	-0.80	0.20	0.62	0.22	31	0.57	0.19	34
		Standardized ELA pre-test score	1.12	0.28	-0.57	0.46	31	-0.45	0.44	34
		Standardized math pre-test score	1.28	0.32	-0.57	0.46	31	-0.43	0.46	34
		Male	-0.16	0.04	0.52	0.08	31	0.52	0.06	34
		White non-Hispanic	1.16	0.29	0.29	0.33	31	0.39	0.33	34
		Black non-Hispanic	-1.00	0.25	0.46	0.37	31	0.38	0.32	34
		Other race	0.05	0.01	0.06	0.06	31	0.06	0.06	34
		Hispanic	-0.17	0.04	0.19	0.25	31	0.18	0.18	34
		Free or reduced-price lunch eligibility	0.05	0.01	0.73	0.26	31	0.74	0.25	34
		English-language learner	-0.45	0.11	0.06	0.07	31	0.05	0.06	34
		Special education status	-2.00**	0.49	0.13	0.09	31	0.10	0.07	34
Multiple types and purposes of writing	Engagement in multiple types of writing	Engagement in multiple types of writing	-2.07**	0.50	0.74	0.35	31	0.55	0.39	34
		Standardized ELA pre-test score	1.12	0.28	-0.57	0.46	31	-0.45	0.44	34
		Standardized math pre-test score	1.28	0.32	-0.57	0.46	31	-0.43	0.46	34
		Male	-0.16	0.04	0.52	0.08	31	0.52	0.06	34
		White non-Hispanic	1.16	0.29	0.29	0.33	31	0.39	0.33	34
		Black non-Hispanic	-1.00	0.25	0.46	0.37	31	0.38	0.32	34
		Other race	0.05	0.01	0.06	0.06	31	0.06	0.06	34
		Hispanic	-0.17	0.04	0.19	0.25	31	0.18	0.18	34
		Free or reduced-price lunch eligibility	0.05	0.01	0.73	0.26	31	0.74	0.25	34
		English-language learner	-0.45	0.11	0.06	0.07	31	0.05	0.06	34
		Special education status	-2.00**	0.49	0.13	0.09	31	0.10	0.07	34

	Construct	Covariate	t-statistic	Effect size	Treatment			Comparison		
					Mean	SD	N	Mean	SD	N
Reading, writing, and/or speaking about texts	Close reading and writing that demonstrates understanding of texts	Close reading and writing that demonstrates understanding of texts	-3.99***	0.89	0.60	0.21	31	0.37	0.26	34
		Standardized ELA pre-test score	1.12	0.28	-0.57	0.46	31	-0.45	0.44	34
		Standardized math pre-test score	1.28	0.32	-0.57	0.46	31	-0.43	0.46	34
		Male	-0.16	0.04	0.52	0.08	31	0.52	0.06	34
		White non-Hispanic	1.16	0.29	0.29	0.33	31	0.39	0.33	34
		Black non-Hispanic	-1.00	0.25	0.46	0.37	31	0.38	0.32	34
		Other race	0.05	0.01	0.06	0.06	31	0.06	0.06	34
		Hispanic	-0.17	0.04	0.19	0.25	31	0.18	0.18	34
		Free or reduced-price lunch eligibility	0.05	0.01	0.73	0.26	31	0.74	0.25	34
		English-language learner	-0.45	0.11	0.06	0.07	31	0.05	0.06	34
		Special education status	-2.00**	0.49	0.13	0.09	31	0.10	0.07	34
Use of text evidence	Text evidence use and argument writing	Text evidence use and argument writing	-3.78***	0.85	0.26	0.24	31	0.08	0.14	34
		Standardized ELA pre-test score	1.12	0.28	-0.57	0.46	31	-0.45	0.44	34
		Standardized math pre-test score	1.28	0.32	-0.57	0.46	31	-0.43	0.46	34
		Male	-0.16	0.04	0.52	0.08	31	0.52	0.06	34
		White non-Hispanic	1.16	0.29	0.29	0.33	31	0.39	0.33	34
		Black non-Hispanic	-1.00	0.25	0.46	0.37	31	0.38	0.32	34
		Other race	0.05	0.01	0.06	0.06	31	0.06	0.06	34
		Hispanic	-0.17	0.04	0.19	0.25	31	0.18	0.18	34
		Free or reduced-price lunch eligibility	0.05	0.01	0.73	0.26	31	0.74	0.25	34
		English-language learner	-0.45	0.11	0.06	0.07	31	0.05	0.06	34
		Special education status	-2.00**	0.49	0.13	0.09	31	0.10	0.07	34
Classroom climate	Classroom climate	Classroom climate	2.35**	0.56	0.64	0.29	31	0.78	0.17	34
		Standardized ELA pre-test score	1.12	0.28	-0.57	0.46	31	-0.45	0.44	34
		Standardized math pre-test score	1.28	0.32	-0.57	0.46	31	-0.43	0.46	34
		Male	-0.16	0.04	0.52	0.08	31	0.52	0.06	34
		White non-Hispanic	1.16	0.29	0.29	0.33	31	0.39	0.33	34
		Black non-Hispanic	-1.00	0.25	0.46	0.37	31	0.38	0.32	34
		Other race	0.05	0.01	0.06	0.06	31	0.06	0.06	34
		Hispanic	-0.17	0.04	0.19	0.25	31	0.18	0.18	34
		Free or reduced-price lunch eligibility	0.05	0.01	0.73	0.26	31	0.74	0.25	34
		English-language learner	-0.45	0.11	0.06	0.07	31	0.05	0.06	34
		Special education status	-2.00**	0.49	0.13	0.09	31	0.10	0.07	34
Classroom management	Observed classroom management	Observed classroom management	-3.23***	0.75	0.52	0.30	31	0.28	0.30	34
		Standardized ELA pre-test score	1.12	0.28	-0.57	0.46	31	-0.45	0.44	34
		Standardized math pre-test score	1.28	0.32	-0.57	0.46	31	-0.43	0.46	34
		Male	-0.16	0.04	0.52	0.08	31	0.52	0.06	34

	Construct	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
Content knowledge		White non-Hispanic	1.16	0.29	0.29	0.33	31	0.39	0.33	34
		Black non-Hispanic	-1.00	0.25	0.46	0.37	31	0.38	0.32	34
		Other race	0.05	0.01	0.06	0.06	31	0.06	0.06	34
		Hispanic	-0.17	0.04	0.19	0.25	31	0.18	0.18	34
		Free or reduced-price lunch eligibility	0.05	0.01	0.73	0.26	31	0.74	0.25	34
		English-language learner	-0.45	0.11	0.06	0.07	31	0.05	0.06	34
		Special education status	-2.00**	0.49	0.13	0.09	31	0.10	0.07	34
	Demonstrate content knowledge in writing	Demonstrate content knowledge in writing	-0.47	0.13	0.82	0.28	28	0.79	0.27	29
		Standardized ELA pre-test score	0.83	0.22	-0.56	0.40	28	-0.46	0.47	29
		Standardized math pre-test score	0.96	0.25	-0.58	0.42	28	-0.46	0.50	29
		Male	0.33	0.09	0.51	0.07	28	0.51	0.05	29
		White non-Hispanic	1.07	0.28	0.30	0.33	28	0.40	0.33	29
		Black non-Hispanic	-0.65	0.17	0.42	0.37	28	0.36	0.32	29
		Other race	-0.33	0.09	0.06	0.06	28	0.06	0.06	29
		Hispanic	-0.48	0.13	0.22	0.27	28	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.29	0.08	0.74	0.26	28	0.76	0.26	29
		English-language learner	-0.95	0.25	0.07	0.10	28	0.05	0.06	29
		Special education status	-2.32**	0.59	0.14	0.07	28	0.09	0.08	29
Higher-order thinking	Assessment of higher order thinking	Assessment of higher order thinking	-1.93*	0.50	0.85	0.22	28	0.72	0.27	29
		Standardized ELA pre-test score	0.83	0.22	-0.56	0.40	28	-0.46	0.47	29
		Standardized math pre-test score	0.96	0.25	-0.58	0.42	28	-0.46	0.50	29
		Male	0.33	0.09	0.51	0.07	28	0.51	0.05	29
		White non-Hispanic	1.07	0.28	0.30	0.33	28	0.40	0.33	29
		Black non-Hispanic	-0.65	0.17	0.42	0.37	28	0.36	0.32	29
		Other race	-0.33	0.09	0.06	0.06	28	0.06	0.06	29
		Hispanic	-0.48	0.13	0.22	0.27	28	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.29	0.08	0.74	0.26	28	0.76	0.26	29
		English-language learner	-0.95	0.25	0.07	0.10	28	0.05	0.06	29
		Special education status	-2.32**	0.59	0.14	0.07	28	0.09	0.08	29
	Higher order thinking and reading	Higher order thinking and reading	0.79	0.21	0.56	0.30	28	0.61	0.24	29
		Standardized ELA pre-test score	0.83	0.22	-0.56	0.40	28	-0.46	0.47	29
		Standardized math pre-test score	0.96	0.25	-0.58	0.42	28	-0.46	0.50	29
		Male	0.33	0.09	0.51	0.07	28	0.51	0.05	29
		White non-Hispanic	1.07	0.28	0.30	0.33	28	0.40	0.33	29
		Black non-Hispanic	-0.65	0.17	0.42	0.37	28	0.36	0.32	29
		Other race	-0.33	0.09	0.06	0.06	28	0.06	0.06	29
		Hispanic	-0.48	0.13	0.22	0.27	28	0.19	0.19	29

	Construct	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
Multimedia use		Free or reduced-price lunch eligibility	0.29	0.08	0.74	0.26	28	0.76	0.26	29
		English-language learner	-0.95	0.25	0.07	0.10	28	0.05	0.06	29
		Special education status	-2.32**	0.59	0.14	0.07	28	0.09	0.08	29
	Online writing	Online writing	-0.51	0.13	0.20	0.34	28	0.16	0.28	29
		Standardized ELA pre-test score	0.83	0.22	-0.56	0.40	28	-0.46	0.47	29
		Standardized math pre-test score	0.96	0.25	-0.58	0.42	28	-0.46	0.50	29
		Male	0.33	0.09	0.51	0.07	28	0.51	0.05	29
		White non-Hispanic	1.07	0.28	0.30	0.33	28	0.40	0.33	29
		Black non-Hispanic	-0.65	0.17	0.42	0.37	28	0.36	0.32	29
		Other race	-0.33	0.09	0.06	0.06	28	0.06	0.06	29
		Hispanic	-0.48	0.13	0.22	0.27	28	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.29	0.08	0.74	0.26	28	0.76	0.26	29
		English-language learner	-0.95	0.25	0.07	0.10	28	0.05	0.06	29
		Special education status	-2.32**	0.59	0.14	0.07	28	0.09	0.08	29
Prior knowledge	Student prior knowledge	Student prior knowledge	0.62	0.16	0.54	0.34	28	0.59	0.32	29
		Standardized ELA pre-test score	0.83	0.22	-0.56	0.40	28	-0.46	0.47	29
		Standardized math pre-test score	0.96	0.25	-0.58	0.42	28	-0.46	0.50	29
		Male	0.33	0.09	0.51	0.07	28	0.51	0.05	29
		White non-Hispanic	1.07	0.28	0.30	0.33	28	0.40	0.33	29
		Black non-Hispanic	-0.65	0.17	0.42	0.37	28	0.36	0.32	29
		Other race	-0.33	0.09	0.06	0.06	28	0.06	0.06	29
		Hispanic	-0.48	0.13	0.22	0.27	28	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.29	0.08	0.74	0.26	28	0.76	0.26	29
		English-language learner	-0.95	0.25	0.07	0.10	28	0.05	0.06	29
		Special education status	-2.32**	0.59	0.14	0.07	28	0.09	0.08	29
		Student prior knowledge	0.39	0.07	0.54	0.28	72	0.56	0.26	55
		Standardized ELA pre-test score	0.40	0.07	-0.51	0.48	72	-0.47	0.52	55
		Standardized math pre-test score	1.32	0.24	-0.56	0.47	72	-0.44	0.56	55
		Male	-0.48	0.09	0.52	0.10	72	0.51	0.07	55
		White non-Hispanic	1.83*	0.33	0.23	0.30	72	0.34	0.35	55
		Black non-Hispanic	-1.60	0.29	0.51	0.34	72	0.42	0.34	55
		Other race	0.21	0.04	0.06	0.07	72	0.06	0.07	55
		Hispanic	-0.31	0.06	0.19	0.22	72	0.18	0.19	55
		Free or reduced-price lunch eligibility	0.99	0.18	0.72	0.28	72	0.77	0.23	55
		English-language learner	0.81	0.15	0.07	0.10	72	0.09	0.21	55
		Special education status	-1.58	0.28	0.13	0.12	72	0.10	0.09	55
Student independence	Student independence	Student independence	0.22	0.06	0.21	0.20	27	0.23	0.20	28
		Standardized ELA pre-test score	0.53	0.14	-0.52	0.44	27	-0.46	0.48	28
		Standardized math pre-test score	0.71	0.19	-0.54	0.47	27	-0.44	0.50	28

	Construct	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
		Male	0.27	0.07	0.51	0.08	27	0.51	0.05	28
		White non-Hispanic	0.81	0.22	0.34	0.34	27	0.41	0.33	28
		Black non-Hispanic	-0.35	0.09	0.37	0.35	27	0.34	0.31	28
		Other race	-0.63	0.17	0.07	0.06	27	0.06	0.06	28
		Hispanic	-0.52	0.14	0.22	0.27	27	0.19	0.19	28
		Free or reduced-price lunch eligibility	0.58	0.16	0.71	0.27	27	0.75	0.26	28
		English-language learner	-0.88	0.24	0.07	0.11	27	0.05	0.06	28
		Special education status	-2.13**	0.56	0.13	0.07	27	0.09	0.08	28
		Student independence	0.91	0.17	0.23	0.16	68	0.25	0.19	55
		Standardized ELA pre-test score	0.14	0.03	-0.48	0.50	68	-0.47	0.52	55
		Standardized math pre-test score	1.04	0.19	-0.54	0.50	68	-0.44	0.56	55
		Male	-0.37	0.07	0.52	0.10	68	0.51	0.07	55
		White non-Hispanic	1.43	0.26	0.25	0.31	68	0.34	0.35	55
		Black non-Hispanic	-1.26	0.23	0.48	0.33	68	0.41	0.34	55
		Other race	-0.02	0.00	0.07	0.07	68	0.06	0.07	55
		Hispanic	-0.23	0.04	0.20	0.22	68	0.19	0.19	55
		Free or reduced-price lunch eligibility	1.21	0.22	0.71	0.29	68	0.77	0.23	55
		English-language learner	0.70	0.13	0.07	0.11	68	0.09	0.21	55
		Special education status	-1.39	0.25	0.13	0.12	68	0.10	0.09	55
Student responsibility for learning	Student responsibility for their learning	Student responsibility for their learning	-0.27	0.07	0.46	0.23	28	0.44	0.16	28
		Standardized ELA pre-test score	0.67	0.18	-0.54	0.44	28	-0.46	0.48	28
		Standardized math pre-test score	0.80	0.22	-0.55	0.47	28	-0.44	0.50	28
		Male	0.27	0.07	0.51	0.07	28	0.51	0.05	28
		White non-Hispanic	0.95	0.25	0.33	0.33	28	0.41	0.33	28
		Black non-Hispanic	-0.56	0.15	0.39	0.36	28	0.34	0.31	28
		Other race	-0.48	0.13	0.07	0.06	28	0.06	0.06	28
		Hispanic	-0.44	0.12	0.22	0.27	28	0.19	0.19	28
		Free or reduced-price lunch eligibility	0.48	0.13	0.72	0.27	28	0.75	0.26	28
		English-language learner	-0.87	0.23	0.07	0.10	28	0.05	0.06	28
		Special education status	-2.21**	0.57	0.13	0.07	28	0.09	0.08	28
		Student responsibility for their learning	-0.60	0.11	0.48	0.20	69	0.46	0.18	55
		Standardized ELA pre-test score	0.22	0.04	-0.49	0.50	69	-0.47	0.52	55
		Standardized math pre-test score	1.09	0.20	-0.54	0.49	69	-0.44	0.56	55
		Male	-0.36	0.07	0.52	0.10	69	0.51	0.07	55
		White non-Hispanic	1.50	0.27	0.25	0.31	69	0.34	0.35	55
		Black non-Hispanic	-1.36	0.24	0.49	0.33	69	0.41	0.34	55
		Other race	0.05	0.01	0.06	0.07	69	0.06	0.07	55
		Hispanic	-0.19	0.03	0.20	0.22	69	0.19	0.19	55

	Construct	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
		Free or reduced-price lunch eligibility	1.15	0.21	0.71	0.29	69	0.77	0.23	55
		English-language learner	0.71	0.13	0.07	0.11	69	0.09	0.21	55
		Special education status	-1.42	0.26	0.13	0.12	69	0.10	0.09	55
Reading and writing instruction										
Academic vocabulary	Academic vocabulary	Academic vocabulary	0.14	0.04	0.38	0.27	31	0.39	0.27	34
		Standardized ELA pre-test score	1.12	0.28	-0.57	0.46	31	-0.45	0.44	34
		Standardized math pre-test score	1.28	0.32	-0.57		31	-0.43	0.46	34
		Male	-0.16	0.04	0.52	0.08	31	0.52	0.06	34
		White non-Hispanic	1.16	0.29	0.29	0.33	31	0.39	0.33	34
		Black non-Hispanic	-1.00	0.25	0.46	0.37	31	0.38	0.32	34
		Other race	0.05	0.01	0.06	0.06	31	0.06	0.06	34
		Hispanic	-0.17	0.04	0.19	0.25	31	0.18	0.18	34
		Free or reduced-price lunch eligibility	0.05	0.01	0.73	0.26	31	0.74	0.25	34
		English-language learner	-0.45	0.11	0.06	0.07	31	0.05	0.06	34
		Special education status	-2.00**	0.49	0.13	0.09	31	0.10	0.07	34
Multiple types and purposes of writing	Writing for multiple purposes and audiences	Writing for multiple purposes and audiences	0.64	0.17	0.82	0.31	28	0.87	0.22	29
		Standardized ELA pre-test score	0.83	0.22	-0.56	0.40	28	-0.46	0.47	29
		Standardized math pre-test score	0.96	0.25	-0.58	0.42	28	-0.46	0.50	29
		Male	0.33	0.09	0.51	0.07	28	0.51	0.05	29
		White non-Hispanic	1.07	0.28	0.30	0.33	28	0.40	0.33	29
		Black non-Hispanic	-0.65	0.17	0.42	0.37	28	0.36	0.32	29
		Other race	-0.33	0.09	0.06	0.06	28	0.06	0.06	29
		Hispanic	-0.48	0.13	0.22	0.27	28	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.29	0.08	0.74	0.26	28	0.76	0.26	29
		English-language learner	-0.95	0.25	0.07	0.10	28	0.05	0.06	29
		Special education status	-2.32**	0.59	0.14	0.07	28	0.09	0.08	29
Reading, writing, and/or speaking about texts	Reading, writing, and/or speaking about texts in past two weeks	Reading, writing, and/or speaking about texts in past two weeks	-1.01	0.27	0.55	0.20	28	0.49	0.22	28
		Standardized ELA pre-test score	0.80	0.21	-0.56	0.40	28	-0.47	0.48	28
		Standardized math pre-test score	0.86	0.23	-0.58	0.42	28	-0.47	0.51	28
		Male	0.38	0.10	0.51	0.07	28	0.51	0.05	28
		White non-Hispanic	1.23	0.33	0.30	0.33	28	0.41	0.33	28
		Black non-Hispanic	-0.89	0.24	0.42	0.37	28	0.34	0.31	28
		Other race	-0.20	0.05	0.06	0.06	28	0.06	0.06	28
		Hispanic	-0.40	0.11	0.22	0.27	28	0.19	0.19	28
		Free or reduced-price lunch eligibility	0.27	0.07	0.74	0.26	28	0.76	0.26	28
		English-language learner	-0.88	0.23	0.07	0.10	28	0.05	0.06	28
		Special education status	-2.32**	0.60	0.14	0.07	28	0.09	0.08	28

	Construct	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
Use of text evidence	Reading, writing, and/or speaking about texts in typical week	Reading, writing, and/or speaking about texts in typical week	-0.38	0.10	0.49	0.27	28	0.46	0.21	29
		Standardized ELA pre-test score	0.83	0.22	-0.56	0.40	28	-0.46	0.47	29
		Standardized math pre-test score	0.96	0.25	-0.58	0.42	28	-0.46	0.50	29
		Male	0.33	0.09	0.51	0.07	28	0.51	0.05	29
		White non-Hispanic	1.07	0.28	0.30	0.33	28	0.40	0.33	29
		Black non-Hispanic	-0.65	0.17	0.42	0.37	28	0.36	0.32	29
		Other race	-0.33	0.09	0.06	0.06	28	0.06	0.06	29
		Hispanic	-0.48	0.13	0.22	0.27	28	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.29	0.08	0.74	0.26	28	0.76	0.26	29
		English-language learner	-0.95	0.25	0.07	0.10	28	0.05	0.06	29
		Special education status	-2.32**	0.59	0.14	0.07	28	0.09	0.08	29
	Assessment of text evidence use	Assessment of text evidence use	-1.58	0.41	0.56	0.13	28	0.50	0.15	29
		Standardized ELA pre-test score	0.83	0.22	-0.56	0.40	28	-0.46	0.47	29
		Standardized math pre-test score	0.96	0.25	-0.58	0.42	28	-0.46	0.50	29
		Male	0.33	0.09	0.51	0.07	28	0.51	0.05	29
		White non-Hispanic	1.07	0.28	0.30	0.33	28	0.40	0.33	29
		Black non-Hispanic	-0.65	0.17	0.42	0.37	28	0.36	0.32	29
		Other race	-0.33	0.09	0.06	0.06	28	0.06	0.06	29
		Hispanic	-0.48	0.13	0.22	0.27	28	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.29	0.08	0.74	0.26	28	0.76	0.26	29
		English-language learner	-0.95	0.25	0.07	0.10	28	0.05	0.06	29
		Special education status	-2.32**	0.59	0.14	0.07	28	0.09	0.08	29
	Self-reported text evidence use	Self-reported text evidence use	-3.85***	0.91	0.72	0.20	28	0.51	0.22	29
		Standardized ELA pre-test score	0.83	0.22	-0.56	0.40	28	-0.46	0.47	29
		Standardized math pre-test score	0.96	0.25	-0.58	0.42	28	-0.46	0.50	29
		Male	0.33	0.09	0.51	0.07	28	0.51	0.05	29
		White non-Hispanic	1.07	0.28	0.30	0.33	28	0.40	0.33	29
		Black non-Hispanic	-0.65	0.17	0.42	0.37	28	0.36	0.32	29
		Other race	-0.33	0.09	0.06	0.06	28	0.06	0.06	29
		Hispanic	-0.48	0.13	0.22	0.27	28	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.29	0.08	0.74	0.26	28	0.76	0.26	29
		English-language learner	-0.95	0.25	0.07	0.10	28	0.05	0.06	29
		Special education status	-2.32**	0.59	0.14	0.07	28	0.09	0.08	29
Writing conventions	Feedback on writing conventions	Feedback on writing conventions	0.73	0.19	0.63	0.30	28	0.68	0.21	29
		Standardized ELA pre-test score	0.83	0.22	-0.56	0.40	28	-0.46	0.47	29
		Standardized math pre-test score	0.96	0.25	-0.58	0.42	28	-0.46	0.50	29
		Male	0.33	0.09	0.51	0.07	28	0.51	0.05	29
		White non-Hispanic	1.07	0.28	0.30	0.33	28	0.40	0.33	29
		Black non-Hispanic	-0.65	0.17	0.42	0.37	28	0.36	0.32	29

	Construct	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
		Other race	-0.33	0.09	0.06	0.06	28	0.06	0.06	29
		Hispanic	-0.48	0.13	0.22	0.27	28	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.29	0.08	0.74	0.26	28	0.76	0.26	29
		English-language learner	-0.95	0.25	0.07	0.10	28	0.05	0.06	29
		Special education status	-2.32**	0.59	0.14	0.07	28	0.09	0.08	29
	Use of writing conventions	Use of writing conventions	3.27***	0.76	0.00	0.00	31	0.21	0.35	34
		Standardized ELA pre-test score	1.12	0.28	-0.57	0.46	31	-0.45	0.44	34
		Standardized math pre-test score	1.28	0.32	-0.57	0.46	31	-0.43	0.46	34
		Male	-0.16	0.04	0.52	0.08	31	0.52	0.06	34
		White non-Hispanic	1.16	0.29	0.29	0.33	31	0.39	0.33	34
		Black non-Hispanic	-1.00	0.25	0.46	0.37	31	0.38	0.32	34
		Other race	0.05	0.01	0.06	0.06	31	0.06	0.06	34
		Hispanic	-0.17	0.04	0.19	0.25	31	0.18	0.18	34
		Free or reduced-price lunch eligibility	0.05	0.01	0.73	0.26	31	0.74	0.25	34
		English-language learner	-0.45	0.11	0.06	0.07	31	0.05	0.06	34
		Special education status	-2.00**	0.49	0.13	0.09	31	0.10	0.07	34
Classroom management and environment										
Classroom management	Classroom disruption	Classroom disruption	1.85*	0.48	0.64	0.30	29	0.77	0.22	29
		Standardized ELA pre-test score	0.55	0.15	-0.53	0.43	29	-0.46	0.47	29
		Standardized math pre-test score	0.61	0.16	-0.54	0.46	29	-0.46	0.50	29
		Male	0.29	0.08	0.51	0.07	29	0.51	0.05	29
		White non-Hispanic	0.93	0.24	0.32	0.33	29	0.40	0.33	29
		Black non-Hispanic	-0.55	0.14	0.41	0.37	29	0.36	0.32	29
		Other race	-0.45	0.12	0.06	0.06	29	0.06	0.06	29
		Hispanic	-0.41	0.11	0.21	0.27	29	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.50	0.13	0.73	0.26	29	0.76	0.26	29
		English-language learner	-0.85	0.22	0.07	0.10	29	0.05	0.06	29
		Special education status	-2.25**	0.57	0.13	0.07	29	0.09	0.08	29
	Self-reported classroom management	Self-reported classroom management	1.22	0.32	0.86	0.25	29	0.93	0.12	29
		Standardized ELA pre-test score	0.55	0.15	-0.53	0.43	29	-0.46	0.47	29
		Standardized math pre-test score	0.61	0.16	-0.54	0.46	29	-0.46	0.50	29
		Male	0.29	0.08	0.51	0.07	29	0.51	0.05	29
		White non-Hispanic	0.93	0.24	0.32	0.33	29	0.40	0.33	29
		Black non-Hispanic	-0.55	0.14	0.41	0.37	29	0.36	0.32	29
		Other race	-0.45	0.12	0.06	0.06	29	0.06	0.06	29
		Hispanic	-0.41	0.11	0.21	0.27	29	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.50	0.13	0.73	0.26	29	0.76	0.26	29
		English-language learner	-0.85	0.22	0.07	0.10	29	0.05	0.06	29

Construct	Covariate	t-statistic	Effect size	Treatment			Comparison		
				Mean	SD	N	Mean	SD	N
	Special education status	-2.25**	0.57	0.13	0.07	29	0.09	0.08	29
Overall ELA instructional practices									
Overall ELA instructional practices	Overall ELA instructional practices	-2.52**	0.60	7.37	1.65	31	6.33	1.67	34
	Standardized ELA pre-test score	1.12	0.28	-0.57	0.46	31	-0.45	0.44	34
	Standardized math pre-test score	1.28	0.32	-0.57	0.46	31	-0.43	0.46	34
	Male	-0.16	0.04	0.52	0.08	31	0.52	0.06	34
	White non-Hispanic	1.16	0.29	0.29	0.33	31	0.39	0.33	34
	Black non-Hispanic	-1.00	0.25	0.46	0.37	31	0.38	0.32	34
	Other race	0.05	0.01	0.06	0.06	31	0.06	0.06	34
	Hispanic	-0.17	0.04	0.19	0.25	31	0.18	0.18	34
	Free or reduced-price lunch eligibility	0.05	0.01	0.73	0.26	31	0.74	0.25	34
	English-language learner	-0.45	0.11	0.06	0.07	31	0.05	0.06	34
	Special education status	-2.00**	0.49	0.13	0.09	31	0.10	0.07	34

Table E.11. Comparison of characteristics of teachers students at baseline, one-year survey samples

Topic	Construct	Covariate	t-statistic	Effect size	Treatment			Comparison		
					Mean	SD	N	Mean	SD	N
General Instruction										
Content knowledge	Demonstrate content knowledge in writing	Novice teachers								
		Demonstrate content knowledge in writing	-0.47	0.13	0.82	0.28	28	0.79	0.27	29
		Standardized ELA pre-test score	0.83	0.22	-0.56	0.40	28	-0.46	0.47	29
		Standardized math pre-test score	0.96	0.25	-0.58	0.42	28	-0.46	0.50	29
		Male	0.33	0.09	0.51	0.07	28	0.51	0.05	29
		White non-Hispanic	1.07	0.28	0.30	0.33	28	0.40	0.33	29
		Black non-Hispanic	-0.65	0.17	0.42	0.37	28	0.36	0.32	29
		Other race	-0.33	0.09	0.06	0.06	28	0.06	0.06	29
		Hispanic	-0.48	0.13	0.22	0.27	28	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.29	0.08	0.74	0.26	28	0.76	0.26	29
		English-language learner	-0.95	0.25	0.07	0.10	28	0.05	0.06	29
		Special education status	-2.32**	0.59	0.14	0.07	28	0.09	0.08	29
		All teachers								
		Demonstrate content knowledge in writing	0.49	0.09	0.78	0.28	72	0.80	0.26	55
		Standardized ELA pre-test score	0.40	0.07	-0.51	0.48	72	-0.47	0.52	55
		Standardized math pre-test score	1.32	0.24	-0.56	0.47	72	-0.44	0.56	55
		Male	-0.48	0.09	0.52	0.10	72	0.51	0.07	55
		White non-Hispanic	1.83*	0.33	0.23	0.30	72	0.34	0.35	55
		Black non-Hispanic	-1.60	0.29	0.51	0.34	72	0.42	0.34	55
		Other race	0.21	0.04	0.06	0.07	72	0.06	0.07	55
		Hispanic	-0.31	0.06	0.19	0.22	72	0.18	0.19	55
		Free or reduced-price lunch eligibility	0.99	0.18	0.72	0.28	72	0.77	0.23	55
		English-language learner	0.81	0.15	0.07	0.10	72	0.09	0.21	55

Topic	Construct	Covariate	t-statistic	Effect size	Treatment			Comparison		
					Mean	SD	N	Mean	SD	N
Higher order thinking		Special education status	-1.58	0.28	0.13	0.12	72	0.10	0.09	55
	Assessment of higher order thinking	Novice teachers								
		Assessment of higher order thinking	-1.93*	0.50	0.85	0.22	28	0.72	0.27	29
		Standardized ELA pre-test score	0.83	0.22	-0.56	0.40	28	-0.46	0.47	29
		Standardized math pre-test score	0.96	0.25	-0.58	0.42	28	-0.46	0.50	29
		Male	0.33	0.09	0.51	0.07	28	0.51	0.05	29
		White non-Hispanic	1.07	0.28	0.30	0.33	28	0.40	0.33	29
		Black non-Hispanic	-0.65	0.17	0.42	0.37	28	0.36	0.32	29
		Other race	-0.33	0.09	0.06	0.06	28	0.06	0.06	29
		Hispanic	-0.48	0.13	0.22	0.27	28	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.29	0.08	0.74	0.26	28	0.76	0.26	29
		English-language learner	-0.95	0.25	0.07	0.10	28	0.05	0.06	29
		Special education status	-2.32**	0.59	0.14	0.07	28	0.09	0.08	29
		All teachers								
		Assessment of higher order thinking	0.88	0.16	0.54	0.25	71	0.58	0.22	55
		Standardized ELA pre-test score	0.46	0.08	-0.51	0.48	71	-0.47	0.52	55
		Standardized math pre-test score	1.41	0.25	-0.57	0.47	71	-0.44	0.56	55
		Male	-0.52	0.09	0.52	0.10	71	0.51	0.07	55
		White non-Hispanic	1.8*	0.32	0.23	0.30	71	0.34	0.35	55
		Black non-Hispanic	-1.62	0.29	0.52	0.34	71	0.42	0.34	55
		Other race	0.48	0.09	0.06	0.07	71	0.06	0.07	55
		Hispanic	-0.32	0.06	0.19	0.22	71	0.18	0.19	55
		Free or reduced-price lunch eligibility	0.98	0.18	0.72	0.28	71	0.77	0.23	55
		English-language learner	0.87	0.16	0.06	0.10	71	0.09	0.21	55
		Special education status	-1.63	0.29	0.13	0.12	71	0.10	0.09	55

Topic	Construct	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
	Higher order thinking and reading	Novice teachers								
		Higher order thinking and reading	0.79	0.21	0.56	0.30	28	0.61	0.24	29
		Standardized ELA pre-test score	0.83	0.22	-0.56	0.40	28	-0.46	0.47	29
		Standardized math pre-test score	0.96	0.25	-0.58	0.42	28	-0.46	0.50	29
		Male	0.33	0.09	0.51	0.07	28	0.51	0.05	29
		White non-Hispanic	1.07	0.28	0.30	0.33	28	0.40	0.33	29
		Black non-Hispanic	-0.65	0.17	0.42	0.37	28	0.36	0.32	29
		Other race	-0.33	0.09	0.06	0.06	28	0.06	0.06	29
		Hispanic	-0.48	0.13	0.22	0.27	28	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.29	0.08	0.74	0.26	28	0.76	0.26	29
		English-language learner	-0.95	0.25	0.07	0.10	28	0.05	0.06	29
		Special education status	-2.32**	0.59	0.14	0.07	28	0.09	0.08	29
		All teachers								
		Higher order thinking and reading	-0.54	0.09	0.60	0.25	71	0.58	0.22	55
		Standardized ELA pre-test score	0.56	0.09	-0.58	0.46	71	-0.54	0.55	55
		Standardized math pre-test score	1.24	0.21	-0.54	0.44	71	-0.44	0.54	55
		Male	-0.24	0.04	0.51	0.09	71	0.51	0.07	55
		White non-Hispanic	1.78*	0.30	0.24	0.30	71	0.33	0.34	55
		Black non-Hispanic	-1.31	0.22	0.50	0.35	71	0.42	0.34	55
		Other race	0.57	0.1	0.06	0.07	71	0.07	0.07	55
		Hispanic	-0.73	0.12	0.21	0.24	71	0.18	0.18	55
		Free or reduced-price lunch eligibility	0.98	0.16	0.71	0.29	71	0.76	0.23	55
		English-language learner	0.71	0.12	0.07	0.10	71	0.09	0.2	55
		Special education status	-0.90	0.15	0.13	0.11	71	0.11	0.1	55
Multimedia use	Online writing	Novice teachers								
		Online writing	-0.51	0.13	0.20	0.34	28	0.16	0.28	29

Topic	Construct	Covariate	t-statistic	Effect size	Treatment			Comparison		
					Mean	SD	N	Mean	SD	N
		Standardized ELA pre-test score	0.83	0.22	-0.56	0.40	28	-0.46	0.47	29
		Standardized math pre-test score	0.96	0.25	-0.58	0.42	28	-0.46	0.50	29
		Male	0.33	0.09	0.51	0.07	28	0.51	0.05	29
		White non-Hispanic	1.07	0.28	0.30	0.33	28	0.40	0.33	29
		Black non-Hispanic	-0.65	0.17	0.42	0.37	28	0.36	0.32	29
		Other race	-0.33	0.09	0.06	0.06	28	0.06	0.06	29
		Hispanic	-0.48	0.13	0.22	0.27	28	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.29	0.08	0.74	0.26	28	0.76	0.26	29
		English-language learner	-0.95	0.25	0.07	0.10	28	0.05	0.06	29
		Special education status	-2.32**	0.59	0.14	0.07	28	0.09	0.08	29
		All teachers								
		Online writing	0.69	0.12	0.14	0.28	70	0.18	0.31	55
		Standardized ELA pre-test score	0.42	0.08	-0.51	0.49	70	-0.47	0.52	55
		Standardized math pre-test score	1.36	0.24	-0.57	0.47	70	-0.44	0.56	55
		Male	-0.52	0.09	0.52	0.10	70	0.51	0.07	55
		White non-Hispanic	1.77*	0.32	0.23	0.30	70	0.34	0.35	55
		Black non-Hispanic	-1.59	0.28	0.51	0.34	70	0.42	0.34	55
		Other race	0.54	0.10	0.06	0.07	70	0.06	0.07	55
		Hispanic	-0.35	0.06	0.19	0.22	70	0.18	0.19	55
		Free or reduced-price lunch eligibility	0.88	0.16	0.73	0.28	70	0.77	0.23	55
		English-language learner	0.86	0.15	0.06	0.10	70	0.09	0.21	55
		Special education status	-1.66*	0.30	0.13	0.12	70	0.10	0.09	55
		Novice teachers								
		Student prior knowledge	0.62	0.16	0.54	0.34	28	0.59	0.32	29
Prior knowledge	Student Prior knowledge	Standardized ELA pre-test score	0.83	0.22	-0.56	0.40	28	-0.46	0.47	29
		Standardized math pre-test score	0.96	0.25	-0.58	0.42	28	-0.46	0.50	29
		Male	0.33	0.09	0.51	0.07	28	0.51	0.05	29

Topic	Construct	Covariate	t-statistic	Effect size	Treatment			Comparison		
					Mean	SD	N	Mean	SD	N
Student independence		White non-Hispanic	1.07	0.28	0.30	0.33	28	0.40	0.33	29
		Black non-Hispanic	-0.65	0.17	0.42	0.37	28	0.36	0.32	29
		Other race	-0.33	0.09	0.06	0.06	28	0.06	0.06	29
		Hispanic	-0.48	0.13	0.22	0.27	28	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.29	0.08	0.74	0.26	28	0.76	0.26	29
		English-language learner	-0.95	0.25	0.07	0.10	28	0.05	0.06	29
		Special education status	-2.32**	0.59	0.14	0.07	28	0.09	0.08	29
		All teachers								
		Student prior knowledge	0.39	0.07	0.54	0.28	72	0.56	0.26	55
		Standardized ELA pre-test score	0.40	0.07	-0.51	0.48	72	-0.47	0.52	55
		Standardized math pre-test score	1.32	0.24	-0.56	0.47	72	-0.44	0.56	55
		Male	-0.48	0.09	0.52	0.10	72	0.51	0.07	55
		White non-Hispanic	1.83*	0.33	0.23	0.30	72	0.34	0.35	55
		Black non-Hispanic	-1.60	0.29	0.51	0.34	72	0.42	0.34	55
		Other race	0.21	0.04	0.06	0.07	72	0.06	0.07	55
		Hispanic	-0.31	0.06	0.19	0.22	72	0.18	0.19	55
		Free or reduced-price lunch eligibility	0.99	0.18	0.72	0.28	72	0.77	0.23	55
		English-language learner	0.81	0.15	0.07	0.10	72	0.09	0.21	55
		Special education status	-1.58	0.28	0.13	0.12	72	0.10	0.09	55
	Student independence	Novice teachers								
		Student independence	0.22	0.06	0.21	0.20	27	0.23	0.20	28
		Standardized ELA pre-test score	0.53	0.14	-0.52	0.44	27	-0.46	0.48	28
		Standardized math pre-test score	0.71	0.19	-0.54	0.47	27	-0.44	0.50	28
		Male	0.27	0.07	0.51	0.08	27	0.51	0.05	28
		White non-Hispanic	0.81	0.22	0.34	0.34	27	0.41	0.33	28
		Black non-Hispanic	-0.35	0.09	0.37	0.35	27	0.34	0.31	28
		Other race	-0.63	0.17	0.07	0.06	27	0.06	0.06	28
		Hispanic	-0.52	0.14	0.22	0.27	27	0.19	0.19	28

Topic	Construct	Covariate	t-statistic	Effect size	Treatment			Comparison		
					Mean	SD	N	Mean	SD	N
Student responsibility for learning		Free or reduced-price lunch eligibility	0.58	0.16	0.71	0.27	27	0.75	0.26	28
		English-language learner	-0.88	0.24	0.07	0.11	27	0.05	0.06	28
		Special education status	-2.13**	0.56	0.13	0.07	27	0.09	0.08	28
		All teachers								
		Student independence	0.91	0.17	0.23	0.16	68	0.25	0.19	55
		Standardized ELA pre-test score	0.14	0.03	-0.48	0.50	68	-0.47	0.52	55
		Standardized math pre-test score	1.04	0.19	-0.54	0.50	68	-0.44	0.56	55
		Male	-0.37	0.07	0.52	0.10	68	0.51	0.07	55
		White non-Hispanic	1.43	0.26	0.25	0.31	68	0.34	0.35	55
		Black non-Hispanic	-1.26	0.23	0.48	0.33	68	0.41	0.34	55
		Other race	-0.02	0.00	0.07	0.07	68	0.06	0.07	55
		Hispanic	-0.23	0.04	0.20	0.22	68	0.19	0.19	55
		Free or reduced-price lunch eligibility	1.21	0.22	0.71	0.29	68	0.77	0.23	55
		English-language learner	0.70	0.13	0.07	0.11	68	0.09	0.21	55
		Special education status	-1.39	0.25	0.13	0.12	68	0.10	0.09	55
	Student responsibility for their learning	Novice teachers								
		Student responsibility for their learning	-0.27	0.07	0.46	0.23	28	0.44	0.16	28
		Standardized ELA pre-test score	0.67	0.18	-0.54	0.44	28	-0.46	0.48	28
		Standardized math pre-test score	0.80	0.22	-0.55	0.47	28	-0.44	0.50	28
		Male	0.27	0.07	0.51	0.07	28	0.51	0.05	28
		White non-Hispanic	0.95	0.25	0.33	0.33	28	0.41	0.33	28
		Black non-Hispanic	-0.56	0.15	0.39	0.36	28	0.34	0.31	28
		Other race	-0.48	0.13	0.07	0.06	28	0.06	0.06	28
		Hispanic	-0.44	0.12	0.22	0.27	28	0.19	0.19	28
		Free or reduced-price lunch eligibility	0.48	0.13	0.72	0.27	28	0.75	0.26	28
		English-language learner	-0.87	0.23	0.07	0.10	28	0.05	0.06	28
		Special education status	-2.21**	0.57	0.13	0.07	28	0.09	0.08	28

Topic	Construct	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
		All teachers								
		Student responsibility for their learning	-0.60	0.11	0.48	0.20	69	0.46	0.18	55
		Standardized ELA pre-test score	0.22	0.04	-0.49	0.50	69	-0.47	0.52	55
		Standardized math pre-test score	1.09	0.20	-0.54	0.49	69	-0.44	0.56	55
		Male	-0.36	0.07	0.52	0.10	69	0.51	0.07	55
		White non-Hispanic	1.50	0.27	0.25	0.31	69	0.34	0.35	55
		Black non-Hispanic	-1.36	0.24	0.49	0.33	69	0.41	0.34	55
		Other race	0.05	0.01	0.06	0.07	69	0.06	0.07	55
		Hispanic	-0.19	0.03	0.20	0.22	69	0.19	0.19	55
		Free or reduced-price lunch eligibility	1.15	0.21	0.71	0.29	69	0.77	0.23	55
		English-language learner	0.71	0.13	0.07	0.11	69	0.09	0.21	55
		Special education status	-1.42	0.26	0.13	0.12	69	0.10	0.09	55
Reading and writing instruction										
		Novice teachers								
		Writing for multiple purposes and audiences	0.64	0.17	0.82	0.31	28	0.87	0.22	29
		Standardized ELA pre-test score	0.83	0.22	-0.56	0.40	28	-0.46	0.47	29
		Standardized math pre-test score	0.96	0.25	-0.58	0.42	28	-0.46	0.50	29
Multiple types and purposes of writing	Writing for multiple purposes and audiences	Male	0.33	0.09	0.51	0.07	28	0.51	0.05	29
		White non-Hispanic	1.07	0.28	0.30	0.33	28	0.40	0.33	29
		Black non-Hispanic	-0.65	0.17	0.42	0.37	28	0.36	0.32	29
		Other race	-0.33	0.09	0.06	0.06	28	0.06	0.06	29
		Hispanic	-0.48	0.13	0.22	0.27	28	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.29	0.08	0.74	0.26	28	0.76	0.26	29
		English-language learner	-0.95	0.25	0.07	0.10	28	0.05	0.06	29
		Special education status	-2.32**	0.59	0.14	0.07	28	0.09	0.08	29
		All teachers								

Topic	Construct	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
Reading, writing, and/or speaking about texts		Writing for multiple purposes and audiences	0.84	0.15	0.83	0.30	72	0.87	0.24	55
		Standardized ELA pre-test score	0.40	0.07	-0.51	0.48	72	-0.47	0.52	55
		Standardized math pre-test score	1.32	0.24	-0.56	0.47	72	-0.44	0.56	55
		Male	-0.48	0.09	0.52	0.10	72	0.51	0.07	55
		White non-Hispanic	1.83*	0.33	0.23	0.30	72	0.34	0.35	55
		Black non-Hispanic	-1.60	0.29	0.51	0.34	72	0.42	0.34	55
		Other race	0.21	0.04	0.06	0.07	72	0.06	0.07	55
		Hispanic	-0.31	0.06	0.19	0.22	72	0.18	0.19	55
		Free or reduced-price lunch eligibility	0.99	0.18	0.72	0.28	72	0.77	0.23	55
		English-language learner	0.81	0.15	0.07	0.10	72	0.09	0.21	55
		Special education status	-1.58	0.28	0.13	0.12	72	0.10	0.09	55
	Reading, writing, and/or speaking about texts in past two weeks	Novice teachers								
		Reading, writing, and/or speaking about texts in past two weeks	-1.01	0.27	0.55	0.20	28	0.49	0.22	28
		Standardized ELA pre-test score	0.80	0.21	-0.56	0.40	28	-0.47	0.48	28
		Standardized math pre-test score	0.86	0.23	-0.58	0.42	28	-0.47	0.51	28
		Male	0.38	0.10	0.51	0.07	28	0.51	0.05	28
		White non-Hispanic	1.23	0.33	0.30	0.33	28	0.41	0.33	28
		Black non-Hispanic	-0.89	0.24	0.42	0.37	28	0.34	0.31	28
		Other race	-0.20	0.05	0.06	0.06	28	0.06	0.06	28
		Hispanic	-0.40	0.11	0.22	0.27	28	0.19	0.19	28
		Free or reduced-price lunch eligibility	0.27	0.07	0.74	0.26	28	0.76	0.26	28
		English-language learner	-0.88	0.23	0.07	0.10	28	0.05	0.06	28
		Special education status	-2.32**	0.60	0.14	0.07	28	0.09	0.08	28
		All teachers								
		Reading, writing, and/or speaking about texts in past two weeks	-0.62	0.11	0.52	0.20	72	0.50	0.21	54

Topic	Construct	Covariate	t-statistic	Effect size	Treatment			Comparison		
					Mean	SD	N	Mean	SD	N
		Standardized ELA pre-test score	0.38	0.07	-0.51	0.48	72	-0.47	0.53	54
		Standardized math pre-test score	1.25	0.22	-0.56	0.47	72	-0.45	0.57	54
		Male	-0.45	0.08	0.52	0.10	72	0.51	0.07	54
		White non-Hispanic	1.93*	0.34	0.23	0.30	72	0.34	0.35	54
		Black non-Hispanic	-1.77*	0.32	0.51	0.34	72	0.41	0.34	54
		Other race	0.30	0.05	0.06	0.07	72	0.07	0.07	54
		Hispanic	-0.24	0.04	0.19	0.22	72	0.18	0.19	54
		Free or reduced-price lunch eligibility	0.97	0.17	0.72	0.28	72	0.77	0.23	54
		English-language learner	0.85	0.15	0.07	0.10	72	0.09	0.21	54
		Special education status	-1.57	0.28	0.13	0.12	72	0.10	0.09	54
	Reading, writing, and/or speaking about texts in typical week	Novice teachers								
		Reading, writing, and/or speaking about texts in typical week	-0.38	0.10	0.49	0.27	28	0.46	0.21	29
		Standardized ELA pre-test score	0.83	0.22	-0.56	0.40	28	-0.46	0.47	29
		Standardized math pre-test score	0.96	0.25	-0.58	0.42	28	-0.46	0.50	29
		Male	0.33	0.09	0.51	0.07	28	0.51	0.05	29
		White non-Hispanic	1.07	0.28	0.30	0.33	28	0.40	0.33	29
		Black non-Hispanic	-0.65	0.17	0.42	0.37	28	0.36	0.32	29
		Other race	-0.33	0.09	0.06	0.06	28	0.06	0.06	29
		Hispanic	-0.48	0.13	0.22	0.27	28	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.29	0.08	0.74	0.26	28	0.76	0.26	29
		English-language learner	-0.95	0.25	0.07	0.10	28	0.05	0.06	29
		Special education status	-2.32**	0.59	0.14	0.07	28	0.09	0.08	29
		All teachers								
		Reading, writing, and/or speaking about texts in typical week	-0.57	0.10	0.50	0.23	72	0.48	0.20	55
		Standardized ELA pre-test score	0.40	0.07	-0.51	0.48	72	-0.47	0.52	55

Topic	Construct	Covariate	t-statistic	Effect size	Treatment			Comparison		
					Mean	SD	N	Mean	SD	N
Use of text evidence		Standardized math pre-test score	1.32	0.24	-0.56	0.47	72	-0.44	0.56	55
		Male	-0.48	0.09	0.52	0.10	72	0.51	0.07	55
		White non-Hispanic	1.83*	0.33	0.23	0.30	72	0.34	0.35	55
		Black non-Hispanic	-1.60	0.29	0.51	0.34	72	0.42	0.34	55
		Other race	0.21	0.04	0.06	0.07	72	0.06	0.07	55
		Hispanic	-0.31	0.06	0.19	0.22	72	0.18	0.19	55
		Free or reduced-price lunch eligibility	0.99	0.18	0.72	0.28	72	0.77	0.23	55
		English-language learner	0.81	0.15	0.07	0.10	72	0.09	0.21	55
		Special education status	-1.58	0.28	0.13	0.12	72	0.10	0.09	55
	Assessment of text evidence use	Novice teachers								
		Assessment of text evidence use	-1.58	0.41	0.56	0.13	28	0.50	0.15	29
		Standardized ELA pre-test score	0.83	0.22	-0.56	0.40	28	-0.46	0.47	29
		Standardized math pre-test score	0.96	0.25	-0.58	0.42	28	-0.46	0.50	29
		Male	0.33	0.09	0.51	0.07	28	0.51	0.05	29
		White non-Hispanic	1.07	0.28	0.30	0.33	28	0.40	0.33	29
		Black non-Hispanic	-0.65	0.17	0.42	0.37	28	0.36	0.32	29
		Other race	-0.33	0.09	0.06	0.06	28	0.06	0.06	29
		Hispanic	-0.48	0.13	0.22	0.27	28	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.29	0.08	0.74	0.26	28	0.76	0.26	29
		English-language learner	-0.95	0.25	0.07	0.10	28	0.05	0.06	29
		Special education status	-2.32**	0.59	0.14	0.07	28	0.09	0.08	29
		All teachers								
		Assessment of text evidence use	-0.58	0.10	0.54	0.16	72	0.52	0.15	55
		Standardized ELA pre-test score	0.40	0.07	-0.51	0.48	72	-0.47	0.52	55
		Standardized math pre-test score	1.32	0.24	-0.56	0.47	72	-0.44	0.56	55
		Male	-0.48	0.09	0.52	0.10	72	0.51	0.07	55

Topic	Construct	Covariate	t-statistic	Effect size	Treatment			Comparison		
					Mean	SD	N	Mean	SD	N
		White non-Hispanic	1.83*	0.33	0.23	0.30	72	0.34	0.35	55
		Black non-Hispanic	-1.60	0.29	0.51	0.34	72	0.42	0.34	55
		Other race	0.21	0.04	0.06	0.07	72	0.06	0.07	55
		Hispanic	-0.31	0.06	0.19	0.22	72	0.18	0.19	55
		Free or reduced-price lunch eligibility	0.99	0.18	0.72	0.28	72	0.77	0.23	55
		English-language learner	0.81	0.15	0.07	0.10	72	0.09	0.21	55
		Special education status	-1.58	0.28	0.13	0.12	72	0.10	0.09	55
	Self-reported text evidence use	Novice teachers								
		Self-reported text evidence use	-3.85***	0.91	0.72	0.20	28	0.51	0.22	29
		Standardized ELA pre-test score	0.83	0.22	-0.56	0.40	28	-0.46	0.47	29
		Standardized math pre-test score	0.96	0.25	-0.58	0.42	28	-0.46	0.50	29
		Male	0.33	0.09	0.51	0.07	28	0.51	0.05	29
		White non-Hispanic	1.07	0.28	0.30	0.33	28	0.40	0.33	29
		Black non-Hispanic	-0.65	0.17	0.42	0.37	28	0.36	0.32	29
		Other race	-0.33	0.09	0.06	0.06	28	0.06	0.06	29
		Hispanic	-0.48	0.13	0.22	0.27	28	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.29	0.08	0.74	0.26	28	0.76	0.26	29
		English-language learner	-0.95	0.25	0.07	0.10	28	0.05	0.06	29
		Special education status	-2.32**	0.59	0.14	0.07	28	0.09	0.08	29
		All teachers								
		Self-reported text evidence use	-2.85***	0.50	0.67	0.21	71	0.55	0.26	55
		Standardized ELA pre-test score	0.36	0.07	-0.50	0.49	71	-0.47	0.52	55
		Standardized math pre-test score	1.27	0.23	-0.56	0.47	71	-0.44	0.56	55
		Male	-0.48	0.09	0.52	0.10	71	0.51	0.07	55
		White non-Hispanic	1.80*	0.32	0.23	0.30	71	0.34	0.35	55
		Black non-Hispanic	-1.57	0.28	0.51	0.34	71	0.42	0.34	55
		Other race	0.26	0.05	0.06	0.07	71	0.06	0.07	55

Topic	Construct	Covariate	t-statistic	Effect size	Treatment			Comparison		
					Mean	SD	N	Mean	SD	N
Writing conventions		Hispanic	-0.34	0.06	0.19	0.22	71	0.18	0.19	55
		Free or reduced-price lunch eligibility	0.89	0.16	0.73	0.28	71	0.77	0.23	55
		English-language learner	0.79	0.14	0.07	0.10	71	0.09	0.21	55
		Special education status	-1.61	0.29	0.13	0.12	71	0.10	0.09	55
	Feedback on writing conventions	Novice teachers								
		Feedback on writing conventions	0.73	0.19	0.63	0.30	28	0.68	0.21	29
		Standardized ELA pre-test score	0.83	0.22	-0.56	0.40	28	-0.46	0.47	29
		Standardized math pre-test score	0.96	0.25	-0.58	0.42	28	-0.46	0.50	29
		Male	0.33	0.09	0.51	0.07	28	0.51	0.05	29
		White non-Hispanic	1.07	0.28	0.30	0.33	28	0.40	0.33	29
		Black non-Hispanic	-0.65	0.17	0.42	0.37	28	0.36	0.32	29
		Other race	-0.33	0.09	0.06	0.06	28	0.06	0.06	29
		Hispanic	-0.48	0.13	0.22	0.27	28	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.29	0.08	0.74	0.26	28	0.76	0.26	29
		English-language learner	-0.95	0.25	0.07	0.10	28	0.05	0.06	29
		Special education status	-2.32**	0.59	0.14	0.07	28	0.09	0.08	29
		All teachers								
		Feedback on writing conventions	1.34	0.24	0.64	0.25	72	0.70	0.22	55
		Standardized ELA pre-test score	0.40	0.07	-0.51	0.48	72	-0.47	0.52	55
		Standardized math pre-test score	1.32	0.24	-0.56	0.47	72	-0.44	0.56	55
		Male	-0.48	0.09	0.52	0.10	72	0.51	0.07	55
		White non-Hispanic	1.83*	0.33	0.23	0.30	72	0.34	0.35	55
		Black non-Hispanic	-1.60	0.29	0.51	0.34	72	0.42	0.34	55
		Other race	0.21	0.04	0.06	0.07	72	0.06	0.07	55
		Hispanic	-0.31	0.06	0.19	0.22	72	0.18	0.19	55
		Free or reduced-price lunch eligibility	0.99	0.18	0.72	0.28	72	0.77	0.23	55

Topic	Construct	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
		English-language learner	0.81	0.15	0.07	0.10	72	0.09	0.21	55
		Special education status	-1.58	0.28	0.13	0.12	72	0.10	0.09	55
	Use of writing conventions	Novice teachers								
		Use of writing conventions	3.27***	0.76	0.00	0.00	31	0.21	0.35	34
		Standardized ELA pre-test score	1.12	0.28	-0.57	0.46	31	-0.45	0.44	34
		Standardized math pre-test score	1.28	0.32	-0.57	0.46	31	-0.43	0.46	34
		Male	-0.16	0.04	0.52	0.08	31	0.52	0.06	34
		White non-Hispanic	1.16	0.29	0.29	0.33	31	0.39	0.33	34
		Black non-Hispanic	-1.00	0.25	0.46	0.37	31	0.38	0.32	34
		Other race	0.05	0.01	0.06	0.06	31	0.06	0.06	34
		Hispanic	-0.17	0.04	0.19	0.25	31	0.18	0.18	34
		Free or reduced-price lunch eligibility	0.05	0.01	0.73	0.26	31	0.74	0.25	34
		English-language learner	-0.45	0.11	0.06	0.07	31	0.05	0.06	34
		Special education status	-2.00**	0.49	0.13	0.09	31	0.10	0.07	34
		All teachers								
		Use of writing conventions	3.27***	0.60	0.01	0.07	52	0.16	0.33	59
		Standardized ELA pre-test score	-0.01	0.00	-0.48	0.46	52	-0.48	0.44	59
		Standardized math pre-test score	0.51	0.10	-0.51	0.45	52	-0.46	0.51	59
		Male	-0.06	0.01	0.51	0.13	52	0.51	0.06	59
		White non-Hispanic	1.23	0.23	0.25	0.31	52	0.33	0.35	59
		Black non-Hispanic	-1.19	0.23	0.49	0.36	52	0.41	0.34	59
		Other race	-1.48	0.28	0.08	0.08	52	0.06	0.06	59
		Hispanic	0.54	0.10	0.18	0.22	52	0.20	0.20	59
		Free or reduced-price lunch eligibility	0.78	0.15	0.73	0.27	52	0.77	0.22	59
		English-language learner	-0.99	0.19	0.06	0.09	52	0.05	0.07	59
		Special education status	0.51	0.10	0.11	0.08	52	0.12	0.09	59
Classroom management and environment										
Classroom management	Classroom disruption	Novice teachers								
		Classroom disruption	1.85*	0.48	0.64	0.30	29	0.77	0.22	29

Topic	Construct	Covariate	t-statistic	Effect size	Treatment			Comparison		
					Mean	SD	N	Mean	SD	N
		Standardized ELA pre-test score	0.55	0.15	-0.53	0.43	29	-0.46	0.47	29
		Standardized math pre-test score	0.61	0.16	-0.54	0.46	29	-0.46	0.50	29
		Male	0.29	0.08	0.51	0.07	29	0.51	0.05	29
		White non-Hispanic	0.93	0.24	0.32	0.33	29	0.40	0.33	29
		Black non-Hispanic	-0.55	0.14	0.41	0.37	29	0.36	0.32	29
		Other race	-0.45	0.12	0.06	0.06	29	0.06	0.06	29
		Hispanic	-0.41	0.11	0.21	0.27	29	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.50	0.13	0.73	0.26	29	0.76	0.26	29
		English-language learner	-0.85	0.22	0.07	0.10	29	0.05	0.06	29
		Special education status	-2.25**	0.57	0.13	0.07	29	0.09	0.08	29
		All teachers								
		Classroom disruption	0.56	0.10	0.73	0.28	73	0.76	0.25	55
		Standardized ELA pre-test score	0.27	0.05	-0.50	0.49	73	-0.47	0.52	55
		Standardized math pre-test score	1.14	0.20	-0.55	0.49	73	-0.44	0.56	55
		Male	-0.50	0.09	0.52	0.10	73	0.51	0.07	55
		White non-Hispanic	1.72*	0.30	0.24	0.30	73	0.34	0.35	55
		Black non-Hispanic	-1.52	0.27	0.51	0.34	73	0.42	0.34	55
		Other race	0.15	0.03	0.06	0.07	73	0.06	0.07	55
		Hispanic	-0.27	0.05	0.19	0.21	73	0.18	0.19	55
		Free or reduced-price lunch eligibility	1.11	0.20	0.72	0.28	73	0.77	0.23	55
		English-language learner	0.85	0.15	0.06	0.10	73	0.09	0.21	55
		Special education status	-1.55	0.28	0.13	0.12	73	0.10	0.09	55
	Self-reported classroom management	Novice teachers								
		Self-reported classroom management	1.22	0.32	0.86	0.25	29	0.93	0.12	29
		Standardized ELA pre-test score	0.55	0.15	-0.53	0.43	29	-0.46	0.47	29
		Standardized math pre-test score	0.61	0.16	-0.54	0.46	29	-0.46	0.50	29

Topic	Construct	Covariate	t-statistic	Effect size	Treatment			Comparison		
					Mean	SD	N	Mean	SD	N
		Male	0.29	0.08	0.51	0.07	29	0.51	0.05	29
		White non-Hispanic	0.93	0.24	0.32	0.33	29	0.40	0.33	29
		Black non-Hispanic	-0.55	0.14	0.41	0.37	29	0.36	0.32	29
		Other race	-0.45	0.12	0.06	0.06	29	0.06	0.06	29
		Hispanic	-0.41	0.11	0.21	0.27	29	0.19	0.19	29
		Free or reduced-price lunch eligibility	0.50	0.13	0.73	0.26	29	0.76	0.26	29
		English-language learner	-0.85	0.22	0.07	0.10	29	0.05	0.06	29
		Special education status	-2.25**	0.57	0.13	0.07	29	0.09	0.08	29
		All teachers								
		Self-reported classroom management	-0.16	0.03	0.91	0.20	74	0.91	0.18	55
		Standardized ELA pre-test score	0.25	0.04	-0.49	0.49	74	-0.47	0.52	55
		Standardized math pre-test score	1.11	0.20	-0.54	0.48	74	-0.44	0.56	55
		Male	-0.60	0.11	0.52	0.10	74	0.51	0.07	55
		White non-Hispanic	1.78*	0.32	0.23	0.30	74	0.34	0.35	55
		Black non-Hispanic	-1.62	0.29	0.52	0.34	74	0.42	0.34	55
		Other race	0.22	0.04	0.06	0.07	74	0.06	0.07	55
		Hispanic	-0.20	0.04	0.19	0.21	74	0.18	0.19	55
		Free or reduced-price lunch eligibility	1.03	0.18	0.72	0.28	74	0.77	0.23	55
		English-language learner	0.88	0.16	0.06	0.10	74	0.09	0.21	55
		Special education status	-1.48	0.26	0.13	0.12	74	0.10	0.09	55
Overall ELA instructional practices (survey)										
		Novice teachers								
		Baseline survey composite	0.17	0.05	8.85	2.47	24	8.94	1.40	25
		Standardized ELA pre-test score	0.64	0.19	-0.52	0.40	24	-0.44	0.48	25
Survey Composite		Standardized math pre-test score	0.96	0.28	-0.54	0.41	24	-0.42	0.51	25
		Male	0.29	0.08	0.51	0.08	24	0.52	0.05	25
		White non-Hispanic	1.11	0.32	0.34	0.34	24	0.45	0.33	25
		Black non-Hispanic	-0.64	0.18	0.37	0.34	24	0.31	0.30	25

Topic	Construct	Covariate	t-statistic	Effect size	Treatment			Comparison		
					Mean	SD	N	Mean	SD	N
		Other race	-0.52	0.15	0.07	0.06	24	0.06	0.06	25
		Hispanic	-0.59	0.17	0.22	0.27	24	0.18	0.19	25
		Free or reduced-price lunch eligibility	0.47	0.14	0.71	0.27	24	0.75	0.27	25
		English-language learner	-0.33	0.10	0.06	0.07	24	0.06	0.07	25
		Special education status	-1.92*	0.53	0.14	0.07	24	0.09	0.08	25
		All teachers								
		Baseline survey composite	-0.19	0.04	8.86	2.18	38	8.78	1.27	38
		Standardized ELA pre-test score	0.06	0.01	-0.40	0.43	38	-0.39	0.47	38
		Standardized math pre-test score	0.79	0.18	-0.46	0.43	38	-0.37	0.56	38
		Male	-0.20	0.05	0.51	0.12	38	0.51	0.06	38
		White non-Hispanic	1.54	0.35	0.30	0.33	38	0.42	0.36	38
		Black non-Hispanic	-1.01	0.23	0.42	0.35	38	0.34	0.32	38
		Other race	-1.00	0.23	0.08	0.07	38	0.06	0.07	38
		Hispanic	-0.58	0.13	0.20	0.24	38	0.17	0.19	38
		Free or reduced-price lunch eligibility	0.77	0.18	0.69	0.29	38	0.74	0.25	38
		English-language learner	-0.04	0.01	0.05	0.07	38	0.05	0.07	38
		Special education status	-0.32	0.07	0.11	0.08	38	0.11	0.09	38

Source: Instructional practice covariates data are from teacher surveys completed in fall 2015 for Cohort 2 schools and fall 2016 for Cohort 3 schools. The pre-test score and demographic data are from administrative educator records from 2014–2015 for Cohort 2 and from 2015–2016 for Cohort 3.

*Statistically significant at the 5 percent level, two-tailed test.

**Statistically significant at the 1 percent level, two-tailed test.

*** Statistically significant at the .01 percent level, two-tailed test.

ELA = English language arts; SD = standard deviation.

Table E.12. Comparison of characteristics of teachers' students at baseline, two-year analysis samples

Topic	Construct	Survey	Obser-vation	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
General Instruction												
Connected lessons	Connected lessons		x	**Novice teachers**								
				Connected lessons	-1.58	0.73	0.62	0.18	11	0.51	0.11	7
				Standardized ELA pre-test score	0.84	0.41	-0.29	0.37	11	-0.13	0.44	7
				Standardized math pre-test score	0.46	0.23	-0.27	0.38	11	-0.18	0.41	7
				English-language learner	-0.36	0.18	0.06	0.07	11	0.05	0.03	7
				Special education status	-0.07	0.03	0.11	0.04	11	0.10	0.08	7
				All teachers								
				Connected lessons	-2.09**	0.78	0.62	0.14	17	0.51	0.10	10
				Standardized ELA pre-test score	-0.32	0.13	-0.25	0.39	17	-0.30	0.49	10
				Standardized math pre-test score	-0.74	0.30	-0.25	0.40	17	-0.38	0.47	10
				English-language learner	-0.61	0.25	0.05	0.06	17	0.04	0.03	10
				Special education status	0.29	0.12	0.10	0.05	17	0.10	0.07	10
	Stated learning purpose		x	**Novice teachers**								
				Stated learning purpose	-1.41	0.66	0.77	0.26	11	0.57	0.35	7
				Standardized ELA pre-test score	0.84	0.41	-0.29	0.37	11	-0.13	0.44	7
				Standardized math pre-test score	0.46	0.23	-0.27	0.38	11	-0.18	0.41	7
				English-language learner	-0.36	0.18	0.06	0.07	11	0.05	0.03	7
				Special education status	-0.07	0.03	0.11	0.04	11	0.10	0.08	7
				All teachers								
				Stated learning purpose	-0.62	0.25	0.68	0.30	17	0.60	0.32	10
				Standardized ELA pre-test score	-0.32	0.13	-0.25	0.39	17	-0.30	0.49	10
				Standardized math pre-test score	-0.74	0.30	-0.25	0.40	17	-0.38	0.47	10
				English-language learner	-0.61	0.25	0.05	0.06	17	0.04	0.03	10
				Special education status	0.29	0.12	0.10	0.05	17	0.10	0.07	10
Connections to world	Learning connected to personal experiences or real world		x	**Novice teachers**								
				Learning connected to personal experiences or real world	0.39	0.19	0.15	0.23	11	0.19	0.18	7
				Standardized ELA pre-test score	0.84	0.41	-0.29	0.37	11	-0.13	0.44	7
				Standardized math pre-test score	0.46	0.23	-0.27	0.38	11	-0.18	0.41	7
				English-language learner	-0.36	0.18	0.06	0.07	11	0.05	0.03	7
				Special education status	-0.07	0.03	0.11	0.04	11	0.10	0.08	7
				All teachers								
				Learning connected to personal experiences or real world	0.13	0.05	0.16	0.24	17	0.17	0.18	10

Topic	Construct	Survey	Obser-vation	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
Content knowledge				Standardized ELA pre-test score	-0.32	0.13	-0.25	0.39	17	-0.30	0.49	10
				Standardized math pre-test score	-0.74	0.30	-0.25	0.40	17	-0.38	0.47	10
				English-language learner	-0.61	0.25	0.05	0.06	17	0.04	0.03	10
				Special education status	0.29	0.12	0.10	0.05	17	0.10	0.07	10
	Content knowledge development		x	**Novice teachers**								
				Content knowledge development	0.00	0.00	0.00	0.00	11	0.00	0.00	7
				Standardized ELA pre-test score	0.84	0.41	-0.29	0.37	11	-0.13	0.44	7
				Standardized math pre-test score	0.46	0.23	-0.27	0.38	11	-0.18	0.41	7
				English-language learner	-0.36	0.18	0.06	0.07	11	0.05	0.03	7
				Special education status	-0.07	0.03	0.11	0.04	11	0.10	0.08	7
				All teachers								
				Content knowledge development	-0.76	0.31	0.05	0.19	17	0.00	0.00	10
				Standardized ELA pre-test score	-0.32	0.13	-0.25	0.39	17	-0.30	0.49	10
				Standardized math pre-test score	-0.74	0.30	-0.25	0.40	17	-0.38	0.47	10
				English-language learner	-0.61	0.25	0.05	0.06	17	0.04	0.03	10
				Special education status	0.29	0.12	0.10	0.05	17	0.10	0.07	10
	Demonstrate content knowledge in writing	x		**Novice teachers**								
				Demonstrate content knowledge in writing	-0.15	0.08	0.92	0.19	12	0.90	0.22	5
				Standardized ELA pre-test score	1.58	0.81	-0.28	0.35	12	0.03	0.43	5
				Standardized math pre-test score	0.77	0.42	-0.24	0.38	12	-0.08	0.46	5
				English-language learner	-0.27	0.15	0.06	0.07	12	0.05	0.03	5
				Special education status	-1.01	0.53	0.11	0.04	12	0.09	0.06	5
				All teachers								
				Demonstrate content knowledge in writing	-0.18	0.07	0.83	0.23	24	0.82	0.25	11
				Standardized ELA pre-test score	0.07	0.03	-0.25	0.39	24	-0.24	0.48	11
				Standardized math pre-test score	0.17	0.06	-0.21	0.38	24	-0.18	0.43	11
				English-language learner	-0.56	0.21	0.06	0.10	24	0.05	0.05	11
				Special education status	-0.05	0.02	0.11	0.06	24	0.11	0.09	11
Higher-order thinking	Assessment of higher order thinking	x		**Novice teachers**								
				Assessment of higher order thinking	-1.51	0.77	0.91	0.14	12	0.78	0.23	5
				Standardized ELA pre-test score	1.58	0.81	-0.28	0.35	12	0.03	0.43	5
				Standardized math pre-test score	0.77	0.42	-0.24	0.38	12	-0.08	0.46	5
				English-language learner	-0.27	0.15	0.06	0.07	12	0.05	0.03	5
				Special education status	-1.01	0.53	0.11	0.04	12	0.09	0.06	5
				All teachers								
				Assessment of higher order thinking	-1.16	0.42	0.87	0.17	24	0.80	0.21	11
				Standardized ELA pre-test score	0.07	0.03	-0.25	0.39	24	-0.24	0.48	11

Topic	Construct	Survey	Obser-vation	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
Multimedia use				Standardized math pre-test score	0.17	0.06	-0.21	0.38	24	-0.18	0.43	11
				English-language learner	-0.56	0.21	0.06	0.10	24	0.05	0.05	11
				Special education status	-0.05	0.02	0.11	0.06	24	0.11	0.09	11
	Higher order thinking and reading	x		**Novice teachers**								
				Higher order thinking and reading	-1.29	0.68	0.73	0.25	12	0.56	0.23	5
				Standardized ELA pre-test score	1.58	0.81	-0.28	0.35	12	0.03	0.43	5
				Standardized math pre-test score	0.77	0.42	-0.24	0.38	12	-0.08	0.46	5
				English-language learner	-0.27	0.15	0.06	0.07	12	0.05	0.03	5
				Special education status	-1.01	0.53	0.11	0.04	12	0.09	0.06	5
				All teachers								
				Higher order thinking and reading	-1.59	0.57	0.70	0.22	24	0.57	0.23	11
				Standardized ELA pre-test score	0.07	0.03	-0.25	0.39	24	-0.24	0.48	11
				Standardized math pre-test score	0.17	0.06	-0.21	0.38	24	-0.18	0.43	11
				English-language learner	-0.56	0.21	0.06	0.10	24	0.05	0.05	11
				Special education status	-0.05	0.02	0.11	0.06	24	0.11	0.09	11
	Observed higher order thinking		x	**Novice teachers**								
				Observed higher order thinking	-1.96*	0.88	0.64	0.12	11	0.48	0.22	7
				Standardized ELA pre-test score	0.84	0.41	-0.29	0.37	11	-0.13	0.44	7
				Standardized math pre-test score	0.46	0.23	-0.27	0.38	11	-0.18	0.41	7
				English-language learner	-0.36	0.18	0.06	0.07	11	0.05	0.03	7
				Special education status	-0.07	0.03	0.11	0.04	11	0.10	0.08	7
				All teachers								
				Observed higher order thinking	-0.26	0.11	0.55	0.21	17	0.53	0.20	10
				Standardized ELA pre-test score	-0.32	0.13	-0.25	0.39	17	-0.30	0.49	10
				Standardized math pre-test score	-0.74	0.30	-0.25	0.40	17	-0.38	0.47	10
				English-language learner	-0.61	0.25	0.05	0.06	17	0.04	0.03	10
				Special education status	0.29	0.12	0.10	0.05	17	0.10	0.07	10
	Online writing	x		**Novice teachers**								
				Online writing	-0.94	0.50	0.22	0.36	12	0.07	0.09	5
				Standardized ELA pre-test score	1.58	0.81	-0.28	0.35	12	0.03	0.43	5
				Standardized math pre-test score	0.77	0.42	-0.24	0.38	12	-0.08	0.46	5
				English-language learner	-0.27	0.15	0.06	0.07	12	0.05	0.03	5
				Special education status	-1.01	0.53	0.11	0.04	12	0.09	0.06	5
				All teachers								
				Online writing	-0.11	0.04	0.19	0.34	24	0.18	0.29	11
				Standardized ELA pre-test score	0.07	0.03	-0.25	0.39	24	-0.24	0.48	11
				Standardized math pre-test score	0.17	0.06	-0.21	0.38	24	-0.18	0.43	11

Topic	Construct	Survey	Obser-vation	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
				English-language learner	-0.56	0.21	0.06	0.10	24	0.05	0.05	11
				Special education status	-0.05	0.02	0.11	0.06	24	0.11	0.09	11
	Use of multimedia in instruction		x	**Novice teachers**								
				Use of multimedia in instruction	-0.18	0.09	0.40	0.25	11	0.38	0.12	7
				Standardized ELA pre-test score	0.84	0.41	-0.29	0.37	11	-0.13	0.44	7
				Standardized math pre-test score	0.46	0.23	-0.27	0.38	11	-0.18	0.41	7
				English-language learner	-0.36	0.18	0.06	0.07	11	0.05	0.03	7
				Special education status	-0.07	0.03	0.11	0.04	11	0.10	0.08	7
				All teachers								
				Use of multimedia in instruction	-0.57	0.23	0.42	0.23	17	0.37	0.19	10
				Standardized ELA pre-test score	-0.32	0.13	-0.25	0.39	17	-0.30	0.49	10
				Standardized math pre-test score	-0.74	0.30	-0.25	0.40	17	-0.38	0.47	10
				English-language learner	-0.61	0.25	0.05	0.06	17	0.04	0.03	10
				Special education status	0.29	0.12	0.10	0.05	17	0.10	0.07	10
Prior knowledge	Student prior knowledge	x		**Novice teachers**								
				Student prior knowledge	0.01	0.01	0.66	0.29	12	0.66	0.41	5
				Standardized ELA pre-test score	1.58	0.81	-0.28	0.35	12	0.03	0.43	5
				Standardized math pre-test score	0.77	0.42	-0.24	0.38	12	-0.08	0.46	5
				English-language learner	-0.27	0.15	0.06	0.07	12	0.05	0.03	5
				Special education status	-1.01	0.53	0.11	0.04	12	0.09	0.06	5
				All teachers								
				Student prior knowledge	-0.71	0.26	0.66	0.26	24	0.59	0.28	11
				Standardized ELA pre-test score	0.07	0.03	-0.25	0.39	24	-0.24	0.48	11
				Standardized math pre-test score	0.17	0.06	-0.21	0.38	24	-0.18	0.43	11
				English-language learner	-0.56	0.21	0.06	0.10	24	0.05	0.05	11
				Special education status	-0.05	0.02	0.11	0.06	24	0.11	0.09	11
Student independence	Student independence	x		**Novice teachers**								
				Student independence	-0.78	0.42	0.30	0.26	11	0.20	0.09	5
				Standardized ELA pre-test score	1.39	0.73	-0.22	0.29	11	0.03	0.43	5
				Standardized math pre-test score	0.54	0.30	-0.19	0.34	11	-0.08	0.46	5
				English-language learner	-0.25	0.14	0.06	0.07	11	0.05	0.03	5
				Special education status	-0.82	0.45	0.11	0.04	11	0.09	0.06	5
				All teachers								
				Student independence	-1.32	0.48	0.31	0.22	23	0.21	0.15	11
				Standardized ELA pre-test score	-0.14	0.05	-0.22	0.37	23	-0.24	0.48	11
				Standardized math pre-test score	-0.03	0.01	-0.18	0.36	23	-0.18	0.43	11
				English-language learner	-0.56	0.21	0.06	0.10	23	0.05	0.05	11

Topic	Construct	Survey	Obser-vation	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
Student participation				Special education status	0.02	0.01	0.11	0.06	23	0.11	0.09	11
	Collaborative discussion practices		x	**Novice teachers**								
				Collaborative discussion practices	0.72	0.35	0.57	0.19	11	0.63	0.15	7
				Standardized ELA pre-test score	0.84	0.41	-0.29	0.37	11	-0.13	0.44	7
				Standardized math pre-test score	0.46	0.23	-0.27	0.38	11	-0.18	0.41	7
				English-language learner	-0.36	0.18	0.06	0.07	11	0.05	0.03	7
				Special education status	-0.07	0.03	0.11	0.04	11	0.10	0.08	7
				All teachers								
				Collaborative discussion practices	1.05	0.42	0.49	0.22	17	0.58	0.15	10
				Standardized ELA pre-test score	-0.32	0.13	-0.25	0.39	17	-0.30	0.49	10
				Standardized math pre-test score	-0.74	0.30	-0.25	0.40	17	-0.38	0.47	10
				English-language learner	-0.61	0.25	0.05	0.06	17	0.04	0.03	10
				Special education status	0.29	0.12	0.10	0.05	17	0.10	0.07	10
	Student engagement in class		x	**Novice teachers**								
				Student engagement in class	0.80	0.39	0.53	0.24	11	0.61	0.16	7
				Standardized ELA pre-test score	0.84	0.41	-0.29	0.37	11	-0.13	0.44	7
				Standardized math pre-test score	0.46	0.23	-0.27	0.38	11	-0.18	0.41	7
				English-language learner	-0.36	0.18	0.06	0.07	11	0.05	0.03	7
				Special education status	-0.07	0.03	0.11	0.04	11	0.10	0.08	7
				All teachers								
				Student engagement in class	0.89	0.35	0.56	0.21	17	0.62	0.13	10
				Standardized ELA pre-test score	-0.32	0.13	-0.25	0.39	17	-0.30	0.49	10
				Standardized math pre-test score	-0.74	0.30	-0.25	0.40	17	-0.38	0.47	10
				English-language learner	-0.61	0.25	0.05	0.06	17	0.04	0.03	10
				Special education status	0.29	0.12	0.10	0.05	17	0.10	0.07	10
	Student participation in discussion		x	**Novice teachers**								
				Student participation in discussion	-0.44	0.22	0.71	0.26	11	0.66	0.14	7
				Standardized ELA pre-test score	0.84	0.41	-0.29	0.37	11	-0.13	0.44	7
				Standardized math pre-test score	0.46	0.23	-0.27	0.38	11	-0.18	0.41	7
				English-language learner	-0.36	0.18	0.06	0.07	11	0.05	0.03	7
				Special education status	-0.07	0.03	0.11	0.04	11	0.10	0.08	7
				All teachers								
				Student participation in discussion	-0.06	0.02	0.64	0.27	17	0.63	0.13	10
				Standardized ELA pre-test score	-0.32	0.13	-0.25	0.39	17	-0.30	0.49	10
				Standardized math pre-test score	-0.74	0.30	-0.25	0.40	17	-0.38	0.47	10
				English-language learner	-0.61	0.25	0.05	0.06	17	0.04	0.03	10
				Special education status	0.29	0.12	0.10	0.05	17	0.10	0.07	10

Topic	Construct	Survey	Obser-vation	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
Student responsibility for learning	Student responsibility for their learning	x		**Novice teachers**								
				Student responsibility for their learning	-1.73	0.87	0.63	0.20	12	0.47	0.11	5
				Standardized ELA pre-test score	1.58	0.81	-0.28	0.35	12	0.03	0.43	5
				Standardized math pre-test score	0.77	0.42	-0.24	0.38	12	-0.08	0.46	5
				English-language learner	-0.27	0.15	0.06	0.07	12	0.05	0.03	5
				Special education status	-1.01	0.53	0.11	0.04	12	0.09	0.06	5
				All teachers								
				Student responsibility for their learning	-0.96	0.35	0.59	0.18	24	0.53	0.20	11
				Standardized ELA pre-test score	0.07	0.03	-0.25	0.39	24	-0.24	0.48	11
				Standardized math pre-test score	0.17	0.06	-0.21	0.38	24	-0.18	0.43	11
				English-language learner	-0.56	0.21	0.06	0.10	24	0.05	0.05	11
				Special education status	-0.05	0.02	0.11	0.06	24	0.11	0.09	11
Reading and writing instruction												
Academic vocabulary	Academic vocabulary		x	**Novice teachers**								
				Academic vocabulary	-1.14	0.55	0.47	0.21	11	0.35	0.24	7
				Standardized ELA pre-test score	0.84	0.41	-0.29	0.37	11	-0.13	0.44	7
				Standardized math pre-test score	0.46	0.23	-0.27	0.38	11	-0.18	0.41	7
				English-language learner	-0.36	0.18	0.06	0.07	11	0.05	0.03	7
				Special education status	-0.07	0.03	0.11	0.04	11	0.10	0.08	7
				All teachers								
				Academic vocabulary	-1.29	0.51	0.44	0.24	17	0.32	0.20	10
				Standardized ELA pre-test score	-0.32	0.13	-0.25	0.39	17	-0.30	0.49	10
				Standardized math pre-test score	-0.74	0.30	-0.25	0.40	17	-0.38	0.47	10
				English-language learner	-0.61	0.25	0.05	0.06	17	0.04	0.03	10
				Special education status	0.29	0.12	0.10	0.05	17	0.10	0.07	10
Multiple types and purposes of writing	Engagement in multiple types of writing		x	**Novice teachers**								
				Engagement in multiple types of writing	-1.33	0.63	0.74	0.34	11	0.54	0.30	7
				Standardized ELA pre-test score	0.84	0.41	-0.29	0.37	11	-0.13	0.44	7
				Standardized math pre-test score	0.46	0.23	-0.27	0.38	11	-0.18	0.41	7
				English-language learner	-0.36	0.18	0.06	0.07	11	0.05	0.03	7
				Special education status	-0.07	0.03	0.11	0.04	11	0.10	0.08	7
				All teachers								
				Engagement in multiple types of writing	-0.98	0.39	0.62	0.38	17	0.47	0.38	10
				Standardized ELA pre-test score	-0.32	0.13	-0.25	0.39	17	-0.30	0.49	10
				Standardized math pre-test score	-0.74	0.30	-0.25	0.40	17	-0.38	0.47	10
				English-language learner	-0.61	0.25	0.05	0.06	17	0.04	0.03	10
				Special education status	0.29	0.12	0.10	0.05	17	0.10	0.07	10

Topic	Construct	Survey	Obser-vation	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
Reading, writing, and/or speaking about texts	Writing for multiple purposes and audiences	x		**Novice teachers**								
				Writing for multiple purposes and audiences	0.94	0.50	0.93	0.17	12	1.00	0.00	5
				Standardized ELA pre-test score	1.58	0.81	-0.28	0.35	12	0.03	0.43	5
				Standardized math pre-test score	0.77	0.42	-0.24	0.38	12	-0.08	0.46	5
				English-language learner	-0.27	0.15	0.06	0.07	12	0.05	0.03	5
				Special education status	-1.01	0.53	0.11	0.04	12	0.09	0.06	5
				All teachers								
				Writing for multiple purposes and audiences	0.91	0.33	0.88	0.21	24	0.95	0.17	11
				Standardized ELA pre-test score	0.07	0.03	-0.25	0.39	24	-0.24	0.48	11
				Standardized math pre-test score	0.17	0.06	-0.21	0.38	24	-0.18	0.43	11
				English-language learner	-0.56	0.21	0.06	0.10	24	0.05	0.05	11
				Special education status	-0.05	0.02	0.11	0.06	24	0.11	0.09	11
	Close reading and writing that demonstrates understanding of texts		x	**Novice teachers**								
				Close reading and writing that demonstrates understanding of texts	-2.18**	0.96	0.61	0.21	11	0.39	0.20	7
				Standardized ELA pre-test score	0.84	0.41	-0.29	0.37	11	-0.13	0.44	7
				Standardized math pre-test score	0.46	0.23	-0.27	0.38	11	-0.18	0.41	7
				English-language learner	-0.36	0.18	0.06	0.07	11	0.05	0.03	7
				Special education status	-0.07	0.03	0.11	0.04	11	0.10	0.08	7
				All teachers								
				Close reading and writing that demonstrates understanding of texts	-0.95	0.38	0.50	0.27	17	0.41	0.24	10
				Standardized ELA pre-test score	-0.32	0.13	-0.25	0.39	17	-0.30	0.49	10
				Standardized math pre-test score	-0.74	0.30	-0.25	0.40	17	-0.38	0.47	10
				English-language learner	-0.61	0.25	0.05	0.06	17	0.04	0.03	10
				Special education status	0.29	0.12	0.10	0.05	17	0.10	0.07	10
	Reading, writing, and/or speaking about texts in past two weeks	x		**Novice teachers**								
				Reading, writing, and/or speaking about texts in past two weeks	-2.92**	1.28	0.68	0.15	12	0.44	0.16	5
				Standardized ELA pre-test score	1.58	0.81	-0.28	0.35	12	0.03	0.43	5
				Standardized math pre-test score	0.77	0.42	-0.24	0.38	12	-0.08	0.46	5
				English-language learner	-0.27	0.15	0.06	0.07	12	0.05	0.03	5
				Special education status	-1.01	0.53	0.11	0.04	12	0.09	0.06	5
				All teachers								
				Reading, writing, and/or speaking about texts in past two weeks	-2.99***	0.98	0.64	0.13	24	0.49	0.16	11

Topic	Construct	Survey	Obser-vation	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
Use of text evidence				Standardized ELA pre-test score	0.07	0.03	-0.25	0.39	24	-0.24	0.48	11
				Standardized math pre-test score	0.17	0.06	-0.21	0.38	24	-0.18	0.43	11
				English-language learner	-0.56	0.21	0.06	0.10	24	0.05	0.05	11
				Special education status	-0.05	0.02	0.11	0.06	24	0.11	0.09	11
	Reading, writing, and/or speaking about texts in typical week	x		**Novice teachers**								
				Reading, writing, and/or speaking about texts in typical week	-1.68	0.85	0.58	0.21	12	0.39	0.21	5
				Standardized ELA pre-test score	1.58	0.81	-0.28	0.35	12	0.03	0.43	5
				Standardized math pre-test score	0.77	0.42	-0.24	0.38	12	-0.08	0.46	5
				English-language learner	-0.27	0.15	0.06	0.07	12	0.05	0.03	5
				Special education status	-1.01	0.53	0.11	0.04	12	0.09	0.06	5
				All teachers								
				Reading, writing, and/or speaking about texts in typical week	-1.90*	0.67	0.57	0.20	24	0.44	0.15	11
				Standardized ELA pre-test score	0.07	0.03	-0.25	0.39	24	-0.24	0.48	11
				Standardized math pre-test score	0.17	0.06	-0.21	0.38	24	-0.18	0.43	11
				English-language learner	-0.56	0.21	0.06	0.10	24	0.05	0.05	11
				Special education status	-0.05	0.02	0.11	0.06	24	0.11	0.09	11
	Assessment of text evidence use	x		**Novice teachers**								
				Assessment of text evidence use	-1.98*	0.97	0.59	0.12	12	0.46	0.14	5
				Standardized ELA pre-test score	1.58	0.81	-0.28	0.35	12	0.03	0.43	5
				Standardized math pre-test score	0.77	0.42	-0.24	0.38	12	-0.08	0.46	5
				English-language learner	-0.27	0.15	0.06	0.07	12	0.05	0.03	5
				Special education status	-1.01	0.53	0.11	0.04	12	0.09	0.06	5
				All teachers								
				Assessment of text evidence use	-1.94*	0.68	0.58	0.15	24	0.48	0.16	11
				Standardized ELA pre-test score	0.07	0.03	-0.25	0.39	24	-0.24	0.48	11
				Standardized math pre-test score	0.17	0.06	-0.21	0.38	24	-0.18	0.43	11
				English-language learner	-0.56	0.21	0.06	0.10	24	0.05	0.05	11
				Special education status	-0.05	0.02	0.11	0.06	24	0.11	0.09	11
	Self-reported text evidence use	x		**Novice teachers**								
				Self-reported text evidence use	-2.83**	1.26	0.81	0.16	12	0.55	0.19	5
				Standardized ELA pre-test score	1.58	0.81	-0.28	0.35	12	0.03	0.43	5
				Standardized math pre-test score	0.77	0.42	-0.24	0.38	12	-0.08	0.46	5
				English-language learner	-0.27	0.15	0.06	0.07	12	0.05	0.03	5
				Special education status	-1.01	0.53	0.11	0.04	12	0.09	0.06	5
				All teachers								

Topic	Construct	Survey	Obser-vation	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
Writing conventions				Self-reported text evidence use	-1.73*	0.61	0.76	0.17	24	0.65	0.20	11
				Standardized ELA pre-test score	0.07	0.03	-0.25	0.39	24	-0.24	0.48	11
				Standardized math pre-test score	0.17	0.06	-0.21	0.38	24	-0.18	0.43	11
				English-language learner	-0.56	0.21	0.06	0.10	24	0.05	0.05	11
				Special education status	-0.05	0.02	0.11	0.06	24	0.11	0.09	11
	Text evidence use and argument writing		x	**Novice teachers**								
				Text evidence use and argument writing	-1.25	0.60	0.14	0.17	11	0.05	0.09	7
				Standardized ELA pre-test score	0.84	0.41	-0.29	0.37	11	-0.13	0.44	7
				Standardized math pre-test score	0.46	0.23	-0.27	0.38	11	-0.18	0.41	7
				English-language learner	-0.36	0.18	0.06	0.07	11	0.05	0.03	7
				Special education status	-0.07	0.03	0.11	0.04	11	0.10	0.08	7
				All teachers								
				Text evidence use and argument writing	-0.59	0.24	0.14	0.17	17	0.10	0.14	10
				Standardized ELA pre-test score	-0.32	0.13	-0.25	0.39	17	-0.30	0.49	10
				Standardized math pre-test score	-0.74	0.30	-0.25	0.40	17	-0.38	0.47	10
				English-language learner	-0.61	0.25	0.05	0.06	17	0.04	0.03	10
				Special education status	0.29	0.12	0.10	0.05	17	0.10	0.07	10
	Feedback on writing conventions	x		**Novice teachers**								
				Feedback on writing conventions	-1.00	0.53	0.77	0.25	12	0.65	0.15	5
				Standardized ELA pre-test score	1.58	0.81	-0.28	0.35	12	0.03	0.43	5
				Standardized math pre-test score	0.77	0.42	-0.24	0.38	12	-0.08	0.46	5
				English-language learner	-0.27	0.15	0.06	0.07	12	0.05	0.03	5
				Special education status	-1.01	0.53	0.11	0.04	12	0.09	0.06	5
				All teachers								
				Feedback on writing conventions	-0.52	0.19	0.74	0.23	24	0.70	0.16	11
				Standardized ELA pre-test score	0.07	0.03	-0.25	0.39	24	-0.24	0.48	11
				Standardized math pre-test score	0.17	0.06	-0.21	0.38	24	-0.18	0.43	11
				English-language learner	-0.56	0.21	0.06	0.10	24	0.05	0.05	11
				Special education status	-0.05	0.02	0.11	0.06	24	0.11	0.09	11
	Use of writing conventions	x		**Novice teachers**								
				Use of writing conventions	1.28	0.61	0.00	0.00	11	0.07	0.19	7
				Standardized ELA pre-test score	0.84	0.41	-0.29	0.37	11	-0.13	0.44	7
				Standardized math pre-test score	0.46	0.23	-0.27	0.38	11	-0.18	0.41	7
				English-language learner	-0.36	0.18	0.06	0.07	11	0.05	0.03	7
				Special education status	-0.07	0.03	0.11	0.04	11	0.10	0.08	7
				All teachers								
				Use of writing conventions	1.32	0.52	0.00	0.00	17	0.05	0.16	10

Topic	Construct	Survey	Obser-vation	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
				Standardized ELA pre-test score	-0.32	0.13	-0.25	0.39	17	-0.30	0.49	10
				Standardized math pre-test score	-0.74	0.30	-0.25	0.40	17	-0.38	0.47	10
				English-language learner	-0.61	0.25	0.05	0.06	17	0.04	0.03	10
				Special education status	0.29	0.12	0.10	0.05	17	0.10	0.07	10
Classroom management and environment												
				Novice teachers								
				Classroom climate	0.35	0.17	0.78	0.29	11	0.82	0.10	7
				Standardized ELA pre-test score	0.84	0.41	-0.29	0.37	11	-0.13	0.44	7
				Standardized math pre-test score	0.46	0.23	-0.27	0.38	11	-0.18	0.41	7
				English-language learner	-0.36	0.18	0.06	0.07	11	0.05	0.03	7
Classroom climate	Classroom climate		x	Special education status	-0.07	0.03	0.11	0.04	11	0.10	0.08	7
				All teachers								
				Classroom climate	0.36	0.14	0.81	0.24	17	0.84	0.09	10
				Standardized ELA pre-test score	-0.32	0.13	-0.25	0.39	17	-0.30	0.49	10
				Standardized math pre-test score	-0.74	0.30	-0.25	0.40	17	-0.38	0.47	10
				English-language learner	-0.61	0.25	0.05	0.06	17	0.04	0.03	10
				Special education status	0.29	0.12	0.10	0.05	17	0.10	0.07	10
				Novice teachers								
				Classroom disruption	0.42	0.23	0.77	0.23	12	0.82	0.24	5
				Standardized ELA pre-test score	1.58	0.81	-0.28	0.35	12	0.03	0.43	5
				Standardized math pre-test score	0.77	0.42	-0.24	0.38	12	-0.08	0.46	5
				English-language learner	-0.27	0.15	0.06	0.07	12	0.05	0.03	5
	Classroom disruption	x		Special education status	-1.01	0.53	0.11	0.04	12	0.09	0.06	5
				All teachers								
				Classroom disruption	-0.10	0.04	0.82	0.20	24	0.81	0.22	10
				Standardized ELA pre-test score	0.68	0.26	-0.25	0.39	24	-0.15	0.40	10
				Standardized math pre-test score	0.62	0.23	-0.21	0.38	24	-0.12	0.39	10
Classroom management				English-language learner	-0.44	0.17	0.06	0.10	24	0.05	0.05	10
				Special education status	-0.04	0.01	0.11	0.06	24	0.11	0.10	10
				Novice teachers								
				Observed classroom management	0.11	0.05	0.30	0.22	11	0.32	0.35	7
				Standardized ELA pre-test score	0.84	0.41	-0.29	0.37	11	-0.13	0.44	7
				Standardized math pre-test score	0.46	0.23	-0.27	0.38	11	-0.18	0.41	7
	Observed classroom management		x	English-language learner	-0.36	0.18	0.06	0.07	11	0.05	0.03	7
				Special education status	-0.07	0.03	0.11	0.04	11	0.10	0.08	7
				All teachers								
				Observed classroom management	-0.58	0.24	0.31	0.20	17	0.26	0.32	10

Topic	Construct	Survey	Obser-vation	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
				Standardized ELA pre-test score	-0.32	0.13	-0.25	0.39	17	-0.30	0.49	10
				Standardized math pre-test score	-0.74	0.30	-0.25	0.40	17	-0.38	0.47	10
				English-language learner	-0.61	0.25	0.05	0.06	17	0.04	0.03	10
				Special education status	0.29	0.12	0.10	0.05	17	0.10	0.07	10
	Self-reported classroom management	x		**Novice teachers**								
				Self-reported classroom management	-2.66**	1.20	0.99	0.01	12	0.95	0.05	5
				Standardized ELA pre-test score	1.58	0.81	-0.28	0.35	12	0.03	0.43	5
				Standardized math pre-test score	0.77	0.42	-0.24	0.38	12	-0.08	0.46	5
				English-language learner	-0.27	0.15	0.06	0.07	12	0.05	0.03	5
				Special education status	-1.01	0.53	0.11	0.04	12	0.09	0.06	5
				All teachers								
				Self-reported classroom management	-3.62***	1.13	0.99	0.01	24	0.96	0.04	11
				Standardized ELA pre-test score	0.07	0.03	-0.25	0.39	24	-0.24	0.48	11
				Standardized math pre-test score	0.17	0.06	-0.21	0.38	24	-0.18	0.43	11
				English-language learner	-0.56	0.21	0.06	0.10	24	0.05	0.05	11
				Special education status	-0.05	0.02	0.11	0.06	24	0.11	0.09	11
Overall ELA instructional practices												
Full composite score		x	x	**Novice teachers**								
				Full composite score	-2.55**	1.18	18.25	1.98	10	15.64	1.57	5
				Standardized ELA pre-test score	1.35	0.72	-0.23	0.31	10	0.03	0.43	5
				Standardized math pre-test score	0.62	0.35	-0.21	0.36	10	-0.08	0.46	5
				English-language learner	-0.42	0.23	0.06	0.08	10	0.05	0.03	5
				Special education status	-0.64	0.36	0.10	0.04	10	0.09	0.06	5
				All teachers								
				Full composite score	-1.51	0.71	17.25	2.51	15	15.60	1.41	6
				Standardized ELA pre-test score	0.72	0.35	-0.18	0.35	15	-0.05	0.43	6
				Standardized math pre-test score	0.02	0.01	-0.18	0.37	15	-0.17	0.47	6
				English-language learner	-0.36	0.18	0.05	0.07	15	0.04	0.03	6
				Special education status	0.24	0.12	0.09	0.05	15	0.10	0.06	6
Survey composite score		x		**Novice teachers**								
				Survey composite score	-1.79*	0.91	10.51	1.71	10	8.99	1.09	5
				Standardized ELA pre-test score	1.35	0.72	-0.23	0.31	10	0.03	0.43	5
				Standardized math pre-test score	0.62	0.35	-0.21	0.36	10	-0.08	0.46	5
				English-language learner	-0.42	0.23	0.06	0.08	10	0.05	0.03	5
				Special education status	-0.64	0.36	0.10	0.04	10	0.09	0.06	5
				All teachers								
				Survey composite score	-1.53	0.71	10.18	1.58	15	9.11	1.01	6

Topic	Construct	Survey	Obser-vation	Covariate	t-statistic	Effect size	Treatment Mean	Treatment SD	Treatment N	Comparison Mean	Comparison SD	Comparison N
Observation composite score				Standardized ELA pre-test score	0.72	0.35	-0.18	0.35	15	-0.05	0.43	6
				Standardized math pre-test score	0.02	0.01	-0.18	0.37	15	-0.17	0.47	6
				English-language learner	-0.36	0.18	0.05	0.07	15	0.04	0.03	6
				Special education status	0.24	0.12	0.09	0.05	15	0.10	0.06	6
			x	**Novice teachers**								
				Observation composite score	-1.38	0.65	7.44	1.49	11	6.49	1.28	7
				Standardized ELA pre-test score	0.84	0.41	-0.29	0.37	11	-0.13	0.44	7
				Standardized math pre-test score	0.46	0.23	-0.27	0.38	11	-0.18	0.41	7
				English-language learner	-0.36	0.18	0.06	0.07	11	0.05	0.03	7
				Special education status	-0.07	0.03	0.11	0.04	11	0.10	0.08	7
				All teachers								
				Observation composite score	-0.96	0.38	6.98	1.67	17	6.40	1.19	10
				Standardized ELA pre-test score	-0.32	0.13	-0.25	0.39	17	-0.30	0.49	10
				Standardized math pre-test score	-0.74	0.30	-0.25	0.40	17	-0.38	0.47	10
				English-language learner	-0.61	0.25	0.05	0.06	17	0.04	0.03	10
				Special education status	0.29	0.12	0.10	0.05	17	0.10	0.07	10

Source: Instructional practice covariates data are from classroom observations or teacher surveys completed in fall 2016. The pre-test score and demographic data are from administrative records from 2015–2016.

*Statistically significant at the 5 percent level, two-tailed test.

**Statistically significant at the 1 percent level, two-tailed test.

*** Statistically significant at the .01 percent level, two-tailed test.

ELA = English language arts; SD = standard deviation.

2. Impact model

To measure impacts using the samples for each outcome and composite measure, we estimated the following ordinary least square regression model that accounted for any remaining differences between treatment and comparison teachers' scores on the baseline version of the outcome[29] (construct), as well as their students' characteristics and prior academic performance:

(V6) $$y_i = \alpha + X_i\beta + \delta T_i + \varepsilon_i$$

where y_i is the outcome of interest for teacher *i*; X_i is a vector of demographic controls (see Table E.13), average baseline test scores, and for teacher *i*, as well as baseline measure of the outcome for the QED analyses only; T_i is a binary variable for treatment status, indicating whether student *i* received support from a specialist in a given subject; ε_i is a random error term that reflects the influence of unobserved factors on the outcome; and δ and β are parameters or vectors of parameters to be estimated, with δ representing the impact of intervention of interest. We also controlled for teachers' clustering within schools by using Huber-White's robust standard errors approach.[30]

Table E.13. Summary of demographic controls included in analyses

	Year 1 impacts		Year 2 impacts	
Demographic controls	Analyses based on observation data	Analyses based on survey data	Analyses based on observation data	Analyses based on survey data
Teacher instructional practice of interest at baseline		✓	✓	✓
Standardized ELA pre-test score	✓	✓	✓	✓
Standardized math pre-test score	✓	✓	✓	✓
Male (percent)	✓	✓		
Race (percent)	✓	✓		
White non-Hispanic	✓	✓		
Black non-Hispanic	✓	✓		
Other race	✓	✓		
Hispanic	✓	✓		
Free or reduced-price lunch eligibility (percent)	✓	✓		
English-language learner (percent)	✓	✓	✓	✓
Special education status (percent)	✓	✓	✓	✓

Note: ✓ indicates the demographic control was included in the regression model.

ELA = English language arts.

[29] We controlled for the baseline version of the outcome for the following outcomes and samples: observation outcomes for the one-year novice sample, survey outcomes for the one-year novice and full samples, and observation and survey outcomes for the two-year novice and full samples.

[30] Due to small sample sizes within a relatively large number of clusters (schools) in the two-year analyses, as well as minimal variation among control variables between groups, we limited the covariates to the following: special education status, English-language learner status, standardized ELA pre-test score, standardized math pre-test score, and the baseline version of the outcome measure.

After each regression model, we computed regression-adjusted mean scores for the outcome for treatment and control teachers and computed effect size (Hedges' g) based on the differences in the regression-adjusted mean scores between treatment and control groups divided by the observed pooled standard deviation (Tables E.14–E.17).

To compensate for the number of inferences being made about instructional practices, we applied the Benjamini-Hochberg multiple comparisons correction to the *p*-values of each pair of impact estimates by subject (Benjamini and Hochberg 1995). Multiple comparisons corrections were conducted for the results of outcomes for each research question; for example, for research question 1, we corrected the *p*-values for the results of all outcomes from the impact analyses of one-year impacts on novice teachers. Significance levels for adjusted *p*-values are provided in the tables. These tables include impact estimates for all outcomes analyzed; a discussion of the statistically significant findings can be found in Chapter V.

Table E.14. All one-year impacts on constructs of novice English language arts teachers' instructional practices

Topic	Outcome	Data source		Estimated treatment coefficient	Standard error	Effect size	p-value	p-value rank	Adjusted critical value	Significant results?	Analytic sample size	
		Obser-vation	Survey								Treatment	Control
General instruction												
Connected lessons	Connected lessons	x		0.22*	0.07	1.27	0.00	2	0.003	Yes	31	34
	Stated learning purpose[a]	x		0.33*	0.10	2.92	0.00	3	0.005	Yes	31	34
Connections to world	Learning connected to personal experiences or real world	x		0.13	0.07	1.34	0.06	8	0.012	No	31	34
Content knowledge	Content knowledge development	x		0.24	0.09	1.40	0.01	4	0.006	No	31	34
	Demonstrate content knowledge in writing		x	0.02	0.08	0.05	0.83	29	0.044	No	28	29
Higher order thinking	Assessment of higher order thinking		x	0.06	0.06	0.72	0.31	17	0.026	No	28	29
	Higher order thinking and reading		x	0.01	0.07	-0.26	0.83	30	0.045	No	28	29
	Observed higher order thinking	x		0.17	0.07	1.01	0.01	5	0.008	No	31	34
Multimedia use	Online writing		x	0.03	0.08	0.23	0.73	26	0.039	No	28	29
	Use of multimedia in instruction	x		0.14	0.08	0.66	0.10	12	0.018	No	31	34
Prior knowledge	Prior knowledge		x	0.00	0.06	-0.14	0.98	33	0.050	No	28	29
Student independence	Student independence		x	-0.01	0.06	0.08	0.92	31	0.047	No	27	28
Student participation	Collaborative discussion practices	x		0.14	0.05	0.89	0.01	6	0.009	No	31	34
	Student engagement in class	x		-0.04	0.09	-0.95	0.61	22	0.033	No	31	34
	Student participation in discussion	x		0.17	0.07	0.61	0.02	7	0.011	No	31	34
Student responsibility for learning	Student responsibility for their learning		x	-0.11	0.06	-0.64	0.09	11	0.017	No	28	28
Reading and writing instruction												
Academic vocabulary	Academic vocabulary	x		0.11	0.07	1.12	0.15	16	0.024	No	31	34
Multiple types and purposes of writing	Engagement in multiple types of writing	x		0.10	0.11	1.03	0.35	18	0.027	No	31	34
	Writing for multiple purposes and audiences		x	-0.05	0.08	-0.08	0.57	21	0.032	No	28	29
	Close reading and writing that demonstrates understanding of texts	x		0.11	0.08	1.55	0.15	15	0.023	No	31	34

Topic	Outcome	Data source		Estimated treatment coefficient	Standard error	Effect size	p-value	p-value rank	Adjusted critical value	Significant results?	Analytic sample size	
		Obser-vation	Survey								Treatment	Control
Reading, writing, and/or speaking about texts	Reading, writing, and/or speaking about texts in past two weeks		x	-0.04	0.06	-0.12	0.54	20	0.030	No	28	28
	Reading, writing, and/or speaking about texts in typical week		x	0.02	0.05	0.24	0.66	24	0.036	No	28	29
Use of text evidence	Assessment of text evidence use		x	0.05	0.06	0.28	0.43	19	0.029	No	28	29
	Self-reported text evidence use		x	0.12	0.07	1.28	0.07	9	0.014	No	28	29
	Text evidence use and argument writing	x		0.13	0.08	2.02	0.09	10	0.015	No	31	34
Writing conventions	Feedback on writing conventions		x	0.01	0.06	0.03	0.93	32	0.048	No	28	29
	Use of writing conventions		x	0.01	0.04	-0.59	0.82	28	0.042	No	31	34
Classroom management and environment												
Classroom climate	Classroom climate	x		0.02	0.05	-0.62	0.65	23	0.035	No	31	34
Classroom management	Classroom disruption		x	-0.10	0.06	-0.75	0.11	14	0.021	No	29	29
	Observed classroom management	x		-0.03	0.08	0.55	0.75	27	0.041	No	31	34
	Self-reported classroom management		x	-0.05	0.03	-0.57	0.10	13	0.020	No	29	29
Overall ELA instructional practices												
Overall ELA instructional practices[a]		x		1.78*	0.47	1.42	0.00	1	0.002	Yes	31	34
			x	-0.17	0.40	-0.18	0.67	25	0.038	No	24	25

Source: For all teachers' observation outcomes, impacts were estimated on teachers' practices in 2015–2016 for Cohort 2 schools and 2016–2017 for Cohort 3 schools. For novice teachers' observation outcomes and novice and all teachers' survey outcomes, impacts were estimated on teachers' practices in spring 2016 for Cohort 2 schools, accounting for fall 2015 as the baseline measure, and spring 2017 for Cohort 3 schools, accounting for fall 2016 as the baseline measure.

Note: Effect size is the adjusted difference between treatment and control group average teacher instructional practice scores divided by the unadjusted pooled standard deviation with an adjustment for small sample sizes per What Works Clearinghouse guidance.

For observation outcomes, the following covariates were included in the regression models: standardized ELA pre-test score, standardized math pre-test score, male, race (white non-Hispanic, black non-Hispanic, and other race), Hispanic ethnicity, free or reduced-price lunch eligibility, English-language learner status, and special education status.

For survey outcomes, the following covariates were included in the regression models: teacher instructional practice of interest at baseline, standardized ELA pre-test score, standardized math pre-test score, male, race (white non-Hispanic, black non-Hispanic, and other race), Hispanic ethnicity, free or reduced-price lunch eligibility, English-language learner status, and special education status.

*Significant differences between control and treatment teachers at the 5 percent level after correcting for multiple comparisons.

[a]Result should be interpreted with caution since there were differences between the groups at baseline on this instructional practice.

ELA = English language arts.

Table E.15. All one-year impacts on constructs of all English language arts teachers' instructional practices

Topic	Outcome	Data source		Estimated treatment coefficient	Standard error	Effect size	p-value	p-value rank	Adjusted critical value	Signi-ficant results?	Analytic sample size	
		Obser-vation	Survey								Treat-ment	Control
General instruction												
Connected lessons	Connected lessons	x		0.07	0.03	0.69	0.03	9	0.014	No	75	89
	Stated learning purpose	x		0.32*	0.05	3.53	0.00	1	0.002	Yes	75	89
Connections to world	Learning connected to personal experiences or real world	x		0.04	0.04	0.57	0.27	17	0.026	No	75	89
Content knowledge	Content knowledge development	x		0.05	0.05	0.50	0.26	16	0.024	No	75	89
	Demonstrate content knowledge in writing		x	0.00	0.05	-0.11	0.94	31	0.047	No	72	55
Higher order thinking	Assessment of higher order thinking		x	0.02	0.04	0.21	0.49	21	0.032	No	72	55
	Higher order thinking and reading		x	0.01	0.04	-0.15	0.80	28	0.042	No	71	55
	Observed higher order thinking	x		0.13*	0.03	1.06	0.00	4	0.006	Yes	75	89
Multimedia use	Online writing		x	-0.01	0.05	-0.12	0.84	29	0.044	No	70	55
	Use of multimedia in instruction	x		0.05	0.03	1.19	0.09	11	0.017	No	75	89
Prior knowledge	Prior knowledge		x	-0.02	0.04	-0.23	0.64	24	0.036	No	72	55
Student independence	Student independence		x	-0.04	0.04	-0.16	0.32	18	0.027	No	68	55
Student participation	Collaborative discussion practices	x		0.12*	0.03	1.24	0.00	3	0.005	Yes	75	89
	Student engagement in class	x		-0.02	0.03	-0.39	0.55	22	0.033	No	75	89
	Student participation in discussion	x		0.08*	0.03	0.79	0.01	7	0.011	Yes	75	89
Student responsibility for learning	Student responsibility for their learning		x	-0.06	0.04	-0.36	0.12	13	0.020	No	69	55
Reading and writing instruction												
Academic vocabulary	Academic vocabulary	x		0.06	0.04	0.79	0.10	12	0.018	No	75	89
Multiple types and purposes of writing	Engagement in multiple types of writing	x		0.02	0.06	0.26	0.67	25	0.038	No	75	89
	Writing for multiple purposes and audiences		x	-0.04	0.05	-0.43	0.48	20	0.030	No	72	55
Reading, writing, and/or speaking about texts	Close reading and writing that demonstrates understanding of texts	x		0.10*	0.04	1.48	0.00	5	0.008	Yes	75	89

Topic	Outcome	Data source		Estimated treatment coefficient	Standard error	Effect size	p-value	p-value rank	Adjusted critical value	Signi-ficant results?	Analytic sample size	
		Obser-vation	Survey								Treat-ment	Control
Use of text evidence	Reading, writing, and/or speaking about texts in past two weeks		x	-0.02	0.04	-0.16	0.56	23	0.035	No	72	54
	Reading, writing, and/or speaking about texts in typical week		x	0.01	0.04	0.20	0.77	26	0.039	No	72	55
	Assessment of text evidence use		x	0.00	0.04	-0.19	0.90	30	0.045	No	72	55
	Self-reported text evidence use[a]		x	0.11*	0.04	0.78	0.01	6	0.009	Yes	71	55
	Text evidence use and argument writing	x		0.09	0.04	1.38	0.02	8	0.012	No	75	89
Writing conventions	Feedback on writing conventions		x	0.00	0.04	-0.12	0.96	32	0.048	No	72	55
	Use of writing conventions		x	0.08	0.04	0.32	0.08	10	0.015	No	52	59
Classroom management and environment												
Classroom climate	Classroom climate	x		0.00	0.03	-0.16	0.99	33	0.050	No	75	89
Classroom management	Classroom disruption		x	-0.05	0.04	-0.31	0.22	15	0.023	No	73	55
	Observed classroom management	x		0.08	0.05	0.62	0.13	14	0.021	No	75	89
	Self-reported classroom management		x	-0.02	0.03	-0.10	0.42	19	0.029	No	74	55
Overall ELA instructional practices												
Overall ELA instructional practices[1]	Observation composite score	x		1.19*	0.24	1.30	0.00	2	0.003	Yes	75	89

Source: For observation outcomes, impacts were estimated on teachers' practices in 2015–2016 for Cohort 2 schools, and 2016–2017 for Cohort 3 schools. For survey outcomes, impacts were estimated on teachers' practices in spring 2016 for Cohort 2 schools, accounting for fall 2015 as the baseline measure, and spring 2017 for Cohort 3 schools, accounting for fall 2016 as the baseline measure.

Note: Effect size is the adjusted difference between treatment and control group average teacher instructional practice scores divided by the unadjusted pooled standard deviation with an adjustment for small sample sizes per What Works Clearinghouse guidance.

For observation outcomes, the following covariates were included in the regression models: standardized ELA pre-test score, standardized math pre-test score, male, race (white non-Hispanic, black non-Hispanic, and other race), Hispanic ethnicity, free or reduced-price lunch eligibility, English-language learner status, and special education status. For survey outcomes, the following covariates were included in the regression models: teacher instructional practice of interest at baseline, standardized ELA pre-test score, standardized math pre-test score, male, race (white non-Hispanic, black non-Hispanic, and other race), Hispanic ethnicity, free or reduced-price lunch eligibility, English-language learner status, and special education status.

*Significant differences between control and treatment teachers at the 5 percent level after correcting for multiple comparisons.
[a]Result should be interpreted with caution since there were differences between the groups at baseline on this instructional practice.
ELA = English language arts.

Table E.16. All two-year impacts on constructs of novice English language arts teachers' instructional practices

Topic	Outcome	Data source		Estimated treatment coefficient	Standard error	Effect size	p-value	p-value rank	Adjusted critical value	Significant results?	Analytic sample size	
		Observation	Survey								Treat-ment	Control
General instruction												
Connected lessons	Connected lessons	x		-0.14	0.14	-0.96	0.35	9	0.014	No	11	7
	Stated learning purpose	x		0.07	0.14	0.76	0.61	17	0.026	No	11	7
Connections to world	Learning connected to personal experiences or real world	x		0.11	0.12	0.34	0.38	11	0.017	No	11	7
Content knowledge	Content knowledge development	x		-0.01	0.08	-0.17	0.91	29	0.044	No	11	7
	Demonstrate content knowledge in writing		x	0.05	0.13	-0.15	0.72	22	0.033	No	12	5
Higher order thinking	Assessment of higher order thinking		x	-0.01	0.10	0.08	0.93	30	0.045	No	12	5
	Higher order thinking and reading		x	-0.05	0.11	0.49	0.65	18	0.027	No	12	5
	Observed higher order thinking	x		0.12	0.13	0.96	0.38	12	0.018	No	11	7
Multimedia use	Online writing		x	-0.03	0.19	0.11	0.88	28	0.042	No	12	5
	Use of multimedia in instruction	x		0.05	0.03	0.75	0.09	3	0.005	No	11	7
Prior knowledge	Prior knowledge		x	0.08	0.08	0.14	0.30	7	0.011	No	12	5
Student independence	Student independence		x	-0.03	0.07	0.18	0.70	20	0.030	No	11	5
Student participation	Collaborative discussion practices	x		0.11	0.10	1.48	0.31	8	0.012	No	11	7
	Student engagement in class	x		-0.02	0.06	-0.30	0.79	25	0.038	No	11	7
	Student participation in discussion	x		0.04	0.11	0.76	0.71	21	0.032	No	11	7
Student responsibility for learning	Student responsibility for their learning		x	-0.03	0.10	0.40	0.76	24	0.036	No	12	5
Reading and writing instruction												
Academic vocabulary	Academic vocabulary	x		0.08	0.13	0.56	0.59	16	0.024	No	11	7
Multiple types and purposes of writing	Engagement in multiple types of writing	x		0.05	0.21	0.12	0.81	26	0.039	No	11	7

Topic	Outcome	Data source		Estimated treatment coefficient	Standard error	Effect size	p-value	p-value rank	Adjusted critical value	Significant results?	Analytic sample size	
		Observation	Survey								Treat-ment	Control
Reading, writing, and/or speaking about texts	Writing for multiple purposes and audiences		x	-0.03	0.05	-0.64	0.51	14	0.021	No	12	5
	Close reading and writing that demonstrates understanding of texts	x		0.00	0.13	0.16	0.99	33	0.050	No	11	7
	Reading, writing, and/or speaking about texts in past two weeks		x	0.18	0.13	1.37	0.17	5	0.008	No	12	5
	Reading, writing, and/or speaking about texts in typical week		x	0.22	0.08	2.04	0.03	2	0.003	No	12	5
Use of text evidence	Assessment of text evidence use		x	0.04	0.13	0.95	0.75	23	0.035	No	12	5
	Self-reported text evidence use		x	0.29	0.11	3.76	0.02	1	0.002	No	12	5
	Text evidence use and argument writing	x		0.13	0.14	0.52	0.37	10	0.015	No	11	7
Writing conventions	Feedback on writing conventions		x	-0.08	0.10	-0.11	0.42	13	0.020	No	12	5
	Use of writing conventions		x	-0.01	0.16	0.21	0.93	32	0.048	No	11	7
Classroom management and environment												
Classroom climate	Classroom climate	x		0.01	0.08	-0.03	0.87	27	0.041	No	11	7
Classroom management	Classroom disruption		x	-0.16	0.10	-0.42	0.14	4	0.006	No	12	5
	Observed classroom management	x		-0.05	0.08	-0.39	0.57	15	0.023	No	11	7
	Self-reported classroom management		x	0.02	0.05	0.56	0.67	19	0.029	No	12	5
Overall ELA instructional practices												
Overall ELA instructional practices		x		0.89	0.74	0.83	0.25	6	0.009	No	11	7
			x	-0.08	0.96	0.63	0.93	31	0.047	No	10	5

Source: Impacts were estimated on teachers' practices in 2017–2018 for Cohort 3 schools, accounting for fall 2016 as the baseline measure.

Note: Effect size is the adjusted difference between treatment and control group average teacher instructional practice scores divided by the unadjusted pooled standard deviation with an adjustment for small sample sizes per What Works Clearinghouse guidance.

The following covariates were included in the regression models: teacher instructional practice of interest at baseline, standardized ELA pre-test score, standardized math pre-test score, English-language learner status, and special education status.

Table E16 (*continued*)

*Significant differences between control and treatment teachers at the 5 percent level after correcting for multiple comparisons.
[a]Result should be interpreted with caution since there were differences between the groups at baseline on this instructional practice.
ELA = English language arts.

Table E.17. All two-year impacts on constructs of all English language arts teachers' instructional practices

Topic	Outcome	Data source		Estimated treatment coefficient	Standard error	Effect size	p-value	p-value rank	Adjusted critical value	Significant results?	Analytic sample size	
		Observation	Survey								Treat-ment	Control
General instruction												
Connected lessons	Connected lessons	x		-0.06	0.10	-0.49	0.58	22	0.033	No	17	10
	Stated learning purpose	x		0.21	0.11	1.95	0.07	6	0.009	No	17	10
Connections to world	Learning connected to personal experiences or real world	x		0.02	0.09	0.10	0.86	30	0.045	No	17	10
Content knowledge	Content knowledge development	x		0.00	0.08	0.31	0.96	33	0.050	No	17	10
	Demonstrate content knowledge in writing		x	-0.04	0.09	-0.31	0.68	25	0.038	No	24	11
Higher order thinking	Assessment of higher order thinking		x	-0.01	0.06	0.05	0.81	28	0.042	No	24	11
	Higher order thinking and reading		x	0.02	0.06	0.68	0.71	26	0.039	No	24	11
	Observed higher order thinking	x		0.17	0.08	1.65	0.05	4	0.006	No	17	10
Multimedia use	Online writing		x	0.08	0.09	0.60	0.36	14	0.021	No	24	11
	Use of multimedia in instruction	x		0.08	0.06	2.74	0.22	10	0.015	No	17	10
Prior knowledge	Prior knowledge		x	-0.02	0.06	0.18	0.74	27	0.041	No	24	11
Student independence	Student independence		x	-0.02	0.04	0.39	0.59	23	0.035	No	23	11
Student participation	Collaborative discussion practices	x		0.20	0.09	3.64	0.04	3	0.005	No	17	10
	Student engagement in class	x		-0.02	0.06	-0.12	0.67	24	0.036	No	17	10
	Student participation in discussion	x		0.13	0.08	1.65	0.12	7	0.011	No	17	10
Student responsibility for learning	Student responsibility for their learning		x	-0.03	0.05	0.05	0.57	21	0.032	No	24	11
Reading and writing instruction												
Academic vocabulary	Academic vocabulary	x		0.11	0.10	1.27	0.31	12	0.018	No	17	10
Multiple types and purposes of writing	Engagement in multiple types of writing	x		0.11	0.17	0.66	0.53	19	0.029	No	17	10
	Writing for multiple purposes and audiences		x	0.10	0.07	1.04	0.14	8	0.012	No	24	11

Topic	Outcome	Data source		Estimated treatment coefficient	Standard error	Effect size	p-value	p-value rank	Adjusted critical value	Significant results?	Analytic sample size	
		Observation	Survey								Treat-ment	Control
Reading, writing, and/or speaking about texts	Close reading and writing that demonstrates understanding of texts	x		0.06	0.08	0.74	0.45	17	0.026	No	17	10
	Reading, writing, and/or speaking about texts in past two weeks		x	0.12	0.06	2.65	0.07	5	0.008	No	24	11
	Reading, writing, and/or speaking about texts in typical week		x	0.06	0.05	1.15	0.25	11	0.017	No	24	11
Use of text evidence	Assessment of text evidence use		x	0.04	0.07	0.64	0.51	18	0.027	No	24	11
	Self-reported text evidence use		x	0.16	0.05	2.16	0.00	1	0.002	No	24	11
	Text evidence use and argument writing	x		0.09	0.11	0.63	0.41	16	0.024	No	17	10
Writing conventions	Feedback on writing conventions		x	-0.01	0.07	0.16	0.93	32	0.048	No	24	11
	Use of writing conventions		x	0.11	0.13	0.78	0.40	15	0.023	No	17	10
Classroom management and environment												
Classroom climate	Classroom climate	x		0.01	0.06	0.30	0.89	31	0.047	No	17	10
Classroom management	Classroom disruption		x	0.10	0.08	1.14	0.20	9	0.014	No	24	10
	Observed classroom management	x		-0.04	0.07	-0.23	0.55	20	0.030	No	17	10
	Self-reported classroom management		x	0.03	0.03	1.06	0.33	13	0.020	No	24	11
Overall ELA instructional practices												
Overall ELA instructional practices		x		1.27	0.55	1.38	0.03	2	0.003	No	17	10
			x	0.15	0.66	0.56	0.82	29	0.044	No	15	6

Source: Impacts were estimated on teachers' practices in 2017–2018 for Cohort 3 schools, accounting for fall 2016 as the baseline measure.

Note: Effect size is the adjusted difference between treatment and control group average teacher instructional practice scores divided by the unadjusted pooled standard deviation with an adjustment for small sample sizes per What Works Clearinghouse guidance.

The following covariates were included in the regression models: teacher instructional practice of interest at baseline, standardized ELA pre-test score, standardized math pre-test score, English-language learner status, and special education status.

*Significant differences between control and treatment teachers at the 5 percent level after correcting for multiple comparisons.

[a]Result should be interpreted with caution since there were differences between the groups at baseline on this instructional practice.

ELA = English language arts.

Appendix F

Earlier TPP study findings

This page has been left blank for double-sided copying.

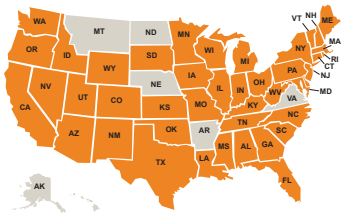
MATHEMATICA
Policy Research

InFOCUS

Jane Choi, Scott Richman, and Sarah Dolfin

Transforming Teachers' Practice: The Impact of EL Education's English Language Arts Curriculum and Professional Learning on Teacher Practices

EL Education



EL Education's literacy curriculum is in use in 44 states plus DC and has been downloaded over 8.7 million times.

ABOUT THE TEACHER POTENTIAL PROJECT

EL Education's TPP features interdisciplinary, content-based English/language arts curriculum alongside engaging professional learning supports for teachers that include (1) five institutes throughout the year; (2) ongoing, personalized, on-site coaching; and (3) online support.

The TPP curriculum and professional learning are closely aligned with the CCSS for English/language arts and literacy, which have been adopted by the majority of states.

The widespread adoption of more rigorous state guidelines for student learning, such as the Common Core State Standards (CCSS), has led to tremendous interest in how best to prepare and develop teachers to foster the skills that students need to meet the new standards and succeed in college and careers. A promising pathway is to combine quality curriculum and professional learning so that teachers' instructional practices are directly aligned with the content they are teaching.[1] While the emerging research suggests the need to integrate teachers' knowledge and skill development with instructional materials, further evidence is necessary to support the effectiveness of such an approach.

EL Education[2] developed the Teacher Potential Project (TPP) to address the growing need to prepare teachers and support student learning through curriculum and embedded professional development. Mathematica Policy Research is conducting an independent, randomized controlled trial evaluation of the TPP to understand its effect on teachers' instructional practices and student achievement. The U.S. Department of Education is funding the project with a five-year Investing in Innovation (i3) grant.

Central to this evaluation and its conceptual framework for student achievement are impacts on teachers' instructional practices after one year of engagement in the TPP. The study team examined a wide range of CCSS-aligned instructional practices and found that TPP-trained teachers engaged students more in reading, writing, and developing content knowledge than teachers who were not trained in the TPP. In addition, the TPP-trained teachers placed more emphasis on having students cite evidence from text, use higher-order thinking skills, and develop responsibility for their own learning.

The study focused on novice teachers,[3] who face the dual challenge of becoming effective educators and meeting the CCSS. However, the study's teacher impact results were similar for both novice and experienced teachers.

[1] Weiner, R. and Pimentel, S. (2017). Practice What You Teach: Connecting Curriculum and Professional Learning in Schools. Washington, D.C.: Aspen Institute.

[2] EL Education is a K-12 educational non-profit organization with the mission of engaging students and teachers in work that is challenging and meaningful so that learning and achievement flourish. EL Education's overarching vision for increasing students' achievement includes three dimensions: mastery of skills and knowledge, character, and high quality student work.

[3] Novice teachers are teachers with 0-3 years of full-time teaching experience, not including substitute and student teaching.

ABOUT THE STUDY

EL Education commissioned Mathematica to conduct a rigorous, independent evaluation of the TPP. The study includes 72 schools in 18 districts across the country, including 10 districts in large, urban areas.

The study team created matched pairs of schools within each district and then randomly assigned schools within those pairs either to adopt the TPP or to continue with the curriculum and professional development offered by the district and school (the control condition). The study team compared teachers' practices across the TPP and control groups by using data from classroom observations conducted by trained members of the research team and surveys of teachers.

As part of its i3 grant evaluation, Mathematica will examine the impact of TPP on student achievement using students' state English/language arts test scores. The study team will estimate one-year and two-year impacts of the TPP on students by comparing those taught by TPP teachers versus those taught by the control teachers. These findings will be available in the summer of 2019.

KEY FINDINGS

The evaluation examined teaching practices that research suggests have the potential for high leverage in preparing students to meet rigorous state learning standards. Impact estimates showed positive effects of the TPP on teachers' CCSS-aligned instructional practices as well as on students' critical thinking skills and engagement with texts.

TPP novice teachers had students do close readings of the text and developed students' content knowledge more often than the control novice teachers (Figure 1). TPP novice teachers also used an average of three writing activities per class period compared to control novice teachers who used an average of two writing activities.

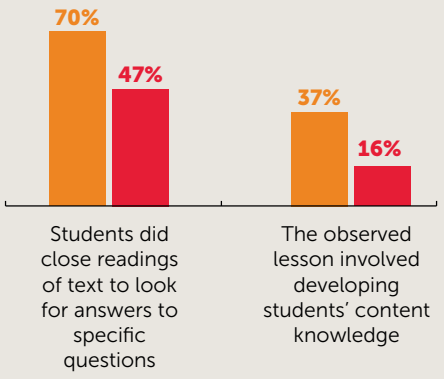


Figure 1

More TPP novice teachers had their students cite evidence from texts, a key research-based skill in the CCSS. A larger proportion of TPP novice teachers had activities focused on understanding the meaning of text compared to the control novice teachers (Figure 2). TPP novice teachers had students reread the text and cite text evidence to support their responses more often than the control novice teachers.

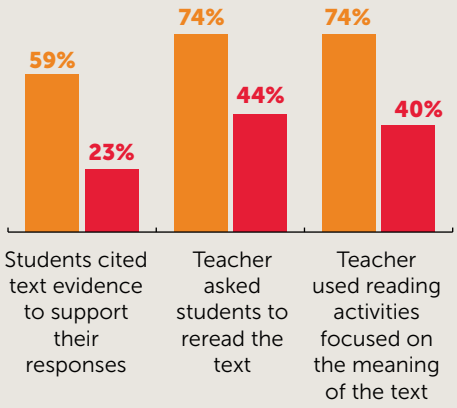


Figure 2

TPP novice teachers more often encouraged students' high-order thinking skills—such as inference, analysis, synthesis, and evaluation. A larger proportion of TPP novice teachers engaged their students' critical thinking skills during classroom discussions compared to the control novice teachers (Figure 3).

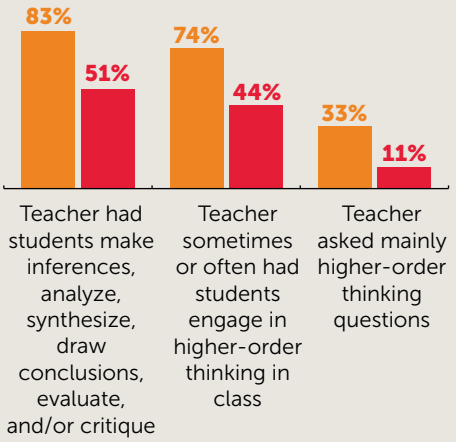


Figure 3

Students taught by TPP novice teachers engaged more often in practices that develop students' responsibility for their own learning. Students did more of the work in developing their content knowledge by collaborating with peers and interacting in class. A larger proportion of students in TPP novice classrooms worked together in pairs or small groups, provided feedback to one another, and talked in class at least as much as the teacher compared to students in the control novice classrooms.

Percentage of TPP novice teachers
Percentage of control novice teachers

For more information, contact Jane Choi at jchoi@mathematica-mpr.com.

Scan this QR code to visit our website.

MATHEMATICA Policy Research

InFOCUS

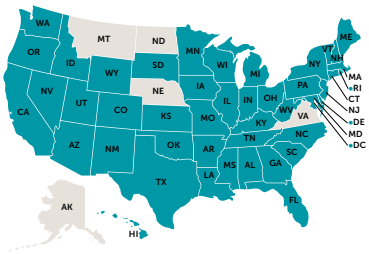
Jane Choi, Scott Richman, and Sarah Dolfin

Sustaining Effective Teacher Practice:
The Impact of the EL Education Language Arts Curriculum and Professional Development on Teachers' Instruction

Teachers play a crucial role in helping students develop the skills needed for success in school, career, and life. These skills—which are a focus in the Common Core State Standards (CCSS)—include higher-order thinking and complex literacy skills, such as reading, writing, and citing evidence from texts.[1] Professional development (PD) can strengthen teachers' instructional practices to help students learn these skills, especially if the PD is paired with curriculum, intensive, sustained over a long duration, and ensures that teachers integrate what they learned.[2]

To address the need for high quality and lasting professional learning for teachers, EL Education developed the Teacher Potential Project (TPP), which includes the CCSS-aligned EL Education Language Arts Curriculum in combination with intensive PD. In an independent study funded by a U.S. Department of Education Investing in Innovation grant, Mathematica Policy Research examined a range of middle grades teachers' CCSS-aligned instructional practices after one and two years of TPP engagement. The study team found that TPP teachers demonstrated and sustained a range of CCSS-aligned instructional practices more than teachers who did not engage with the TPP.[3] These two-year findings build on previous positive results of one year of TPP engagement.[4]

EL Education



The EL Education Language Arts Curriculum is in use in 45 states plus the District of Columbia and has been downloaded more than 10 million times.

ABOUT THE TEACHER POTENTIAL PROJECT

The TPP features the interdisciplinary, content-based EL Education Language Arts Curriculum alongside engaging professional learning supports for teachers that include (1) learning institutes throughout the year; (2) ongoing, personalized on-site coaching; and (3) online support.

KEY FINDINGS

Significantly more TPP teachers encouraged students' higher-order thinking skills—such as inference, analysis, synthesis, and evaluation—in both the first and second years. A larger proportion of TPP teachers engaged their students' critical thinking skills and focused on deeper textual analysis compared with the control teachers (Figure 1).

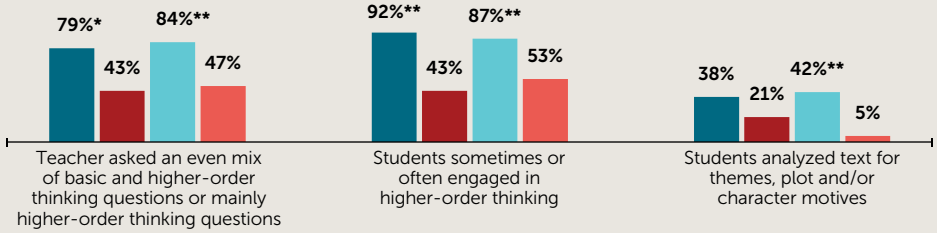


Figure 1

TPP treatment, Year 1 | Control, Year 1 | TPP treatment, Year 2 | Control, Year 2

TPP treatment and control groups significantly different from each other at the .05 (*) or .01 (**) level, two-tailed test.

NOVEMBER 2018

ABOUT THE STUDY

This two-year study included teachers of students in grades 4 through 8 in 18 schools across five districts. Ten of the schools used the EL Education Language Arts Curriculum and PD (TPP treatment) for two years and 8 schools used the curriculum and PD provided by their districts (control).

To learn about the effects of the TPP, the study team compared teachers' practices across the TPP and control groups using data from teacher surveys and classroom observations conducted by the study team.

The TPP and control teachers were similar in their demographic characteristics, average number of years teaching, education background, and certification area at the start of the study. This gives greater confidence that findings from the study are due to the EL Education language arts curriculum and PD and not differences across the teachers in each group.

Significantly more TPP teachers asked students to cite evidence from texts in discussions and their writing. A larger proportion of TPP teachers engaged students in writing activities using text-based evidence in both the first and second years of the project (Figure 2). In addition, in the second year, a greater proportion of TPP teachers prompted students to cite evidence from texts to support their responses verbally.

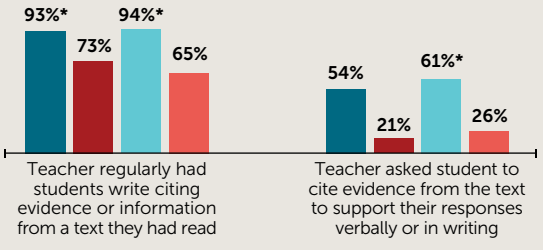


Figure 2

TPP treatment, Year 1 | TPP treatment, Year 2
Control, Year 1 | Control, Year 2

TPP treatment and control groups significantly different from each other at the .05 (*) or .01 (**) level, two-tailed test.

Students taught by TPP teachers engaged significantly more often in reading, writing, and speaking about texts in the second year of TPP. A larger proportion of TPP teachers asked students to collaborate in a group activity or project about what they read, had students explain or support their understanding of what they read, and had students share their ideas or understanding of what they read compared to control teachers. (Figure 3).

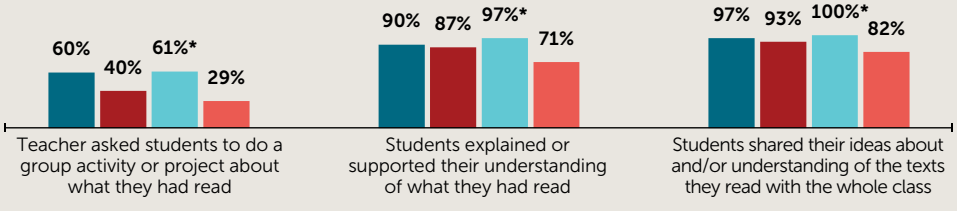


Figure 3

TPP treatment, Year 1 | Control, Year 1 | TPP treatment, Year 2 | Control, Year 2

TPP treatment and control groups significantly different from each other at the .05 (*) or .01 (**) level, two-tailed test.

NEXT STEPS

This study is part of a larger, rigorous randomized controlled trial evaluation in which the study team will examine the impact of the TPP on student achievement using students' state English language arts test scores with the full sample of study schools (72 schools in 18 districts). The study team will estimate one- and two-year impacts of the TPP on student achievement by comparing those taught by TPP teachers with those taught by control teachers. These findings will be available in summer 2019.

ENDNOTES

[1] Common Core State Standards Initiative. "Key Shifts in English Language Arts." CCSS Initiative, 2010. Available at http://www.corestandards.org/other-resources/key-shifts-in-english-language-arts/. Accessed September 5, 2018.

[2] Darling-Hammond, L., M.E. Hyler, and M. Gardner. "Effective Teacher Professional Development." Palo Alto, CA: Learning Policy Institute, 2017.

[3] Analyses included only the teachers who were part of the evaluation in both school years (2016–2017 and 2017–2018) and included both experienced and novice teachers (those with zero to three years of teaching experience in 2016–2017).

[4] Choi, J., S. Richman, and S. Dolfin. "Transforming Teachers' Practice: The Impact of EL Education's English Language Arts Curriculum and Professional Learning on Teacher Practices." Oakland, CA: Mathematica Policy Research, 2017.

For more information, contact Jane Choi at jchoi@mathematica-mpr.com.

Scan this QR code to visit our website.

Appendix G

Supplemental information for student English language arts achievement analyses and student literacy task analyses

This page has been left blank for double-sided copying.

A. Supplemental information for student English language arts achievement analyses

This appendix provides additional information on student ELA achievement analyses in reference to the following four issues discussed in Chapter VI:

- **Attrition of schools and students from the study (Tables G.1 and G.2).** These tables report sample attrition as the number of units (schools and students) in treatment or control conditions during the baseline year and the numbers of units from those groups that remain in the study sample for the first and second years of the study. We report overall and differential attrition, as well as the expected bias from this attrition, using What Works Clearinghouse (WWC) "optimistic" assumptions. Given the many other factors that drive student and staff mobility, we think that the optimistic assumptions are most appropriate for determining whether attrition is sufficiently small to make the impact estimate valid.

- **Selected characteristics of students in analytic samples (Tables G.3 and G.4).** These tables show sociodemographic characteristics of the two selected samples used in analyses of the one-year impact (addressing research question S2) and overall two-year impact (addressing research question S4).

- **Measures of baseline equivalence of student achievement outcomes between treatment and control students for the benchmark model (Tables G.5 and G.6).** These tables report comparisons of baseline equivalence on ELA baseline measures and grade level between students in treatment and control study schools who are included in the final analytic samples. According to the WWC, effect sizes greater than 0.25 standard deviation are interpreted as evidence that the treatment and control groups of students are not sufficiently comparable. Effect sizes of larger than 0.05 and smaller than 0.25 indicate a difference that requires statistical adjustments on the variable. For the dichotomous grade variable (Table G.6), we assess baseline equivalence using absolute percentage point differences between the intervention and comparison groups with a cutoff of 10 percentage points (that corresponds to the WWC benchmark of 0.25 standard deviations).

- **Sensitivity analyses of estimated impacts to alternative specifications (Table G.7).** This table presents additional sensitivity analyses of the estimated impacts on student ELA achievement based on different assumptions and model specifications. Table G.7 summarizes the results of these analyses for four selected impacts (addressing research questions S2, S4, S6, and S9).

Table G.1. School attrition for one-year impact analysis

Analyses and research questions addressed	Overall (N)	Treatment (*N*)	Control (*N*)	Overall/ differential attrition (percent)	WWC attrition rating (using optimistic boundary)
One-year impacts					
Number of schools at random assignment	79	40	39		
Number of schools in sample for one year impact analysis (S2)	61	31	30	22.8/0.6	Low
Number of schools in sample for one year novice impact analysis (S1)	50	24	26	36.7/6.7	Low

Source: Administrative records for 2014–2015 for Cohort 1 pilot schools, 2015–2016 for Cohort 2 schools, and 2016–2017 for Cohort 3 schools.

Note: Number of schools reported. The count of overall sample sizes in this table reflect the original numbers of schools at the time of random assignment. Five schools (3 treatment and 2 control schools) that were each split into two during the study period are subtracted from the total count of 66 schools (in the numerator of attrition formula for the analysis addressing research question S2) in order to provide conservative estimates of attrition rates.

WWC = What Works Clearinghouse.

Table G.2. Student attrition for one-year impact analyses

Analyses and research questions addressed	Overall (*N*)	Treatment (*N*)	Control (*N*)	Overall/ differential attrition (percent)	WWC attrition rating (using optimistic boundary)
(1) One-year impacts (S2)					
Number of students at baseline	15,566	7,061	8,505		
Number remaining at the end of Year 1 *with* ELA score	12,859	6,150	6,709	17.4/8.2	Low
(2) One-year novice impacts (S1)					
Number of students at baseline	5,462	2,432	3,030		
Number remaining at the end of Year 1 *with* ELA score	4,815	2,162	2,653	11.8/1.3	Low

Source: Administrative records for 2014–2015 for Cohort 1 pilot schools, 2015–2016 for Cohort 2 schools, and 2016–2017 for Cohort 3 schools.

Note: Baseline refers to October of the relevant school year.

ELA = English language arts; WWC = What Works Clearinghouse.

Table G.3. Characteristics of student samples in the one-year impact analysis

One-year impact (addressing research question S2)	Treatment group mean (SD)	Control group mean (SD)	Overall sample mean (SD)
Female (percent)	48.5 (0.50)	49.6 (0.50)	49.1 (0.50)
Race/ethnicity (percent)			
White non-Hispanic	31.7 (0.47)	42.7 (0.49)	37.5 (0.48)
Black non-Hispanic	40.7 (0.49)	28.8 (0.45)	34.5 (0.48)
Hispanic	19.6 (0.40)	20.9 (0.41)	20.3 (0.40)
Other	7.7 (0.27)	7.6 (0.26)	7.6 (0.27)
Eligibility for free or reduced-price lunch (percent)	57.9 (0.49)	62.4 (0.48)	60.2 (0.49
English-language learner status (percent)	7.2 (0.26)	8.9 (0.28)	8.0 (0.27)
Special education status (percent)	11.1 (0.31)	10.5 (0.31)	10.8 (0.31)
Sample size (number of students)	6,150	6,709	12,859

SD = standard deviation.

Table G.4. Characteristics of student samples in the overall two-year impact analysis

Overall two-year impact (addressing research question S4)	Treatment group mean (SD)	Control group mean (SD)	Overall sample mean (SD)
Female (percent)	49.2 (0.50)	49.8 (0.50)	49.5 (0.50)
Race/ethnicity (percent)			
White non-Hispanic	29.7 (0.46)	39.1 (0.49)	34.3 (0.47)
Black non-Hispanic	40.1 (0.49)	25.5 (0.44)	32.9 (0.47)
Hispanic	21.7 (0.41)	25.6 (0.44)	23.6 (0.42)
Other	8.2 (0.27)	9.7 (0.30)	8.9 (0.29)
Eligibility for free or reduced-price lunch (percent)	58.2 (0.49)	64.7 (0.48)	61.4 (0.49)
English-language learner status (percent)	6.3 (0.24)	6.3 (0.24)	6.3 (0.24)
Special education status (percent)	12.1 (0.33)	11.0 (0.31)	11.6 (0.32)
Sample size (number of students)	5,168	4,991	10,159

SD = standard deviation.

Table G.5. Comparison of baseline English language arts measures for student impact analyses

	Analyses and research questions addressed	Condition	*N*	Mean	SD	Effect size
One-year impacts in all schools	(1) One-year novice impact (S1)	Treatment	2,162	–0.41	1.00	–0.07
		Control	2,653	–0.34	0.99	
	(2) One-year impact (S2)	Treatment	6,150	–0.34	1.02	0.03
		Control	6,709	–0.37	1.01	
Overall two-year impacts in all schools	(3) Overall two-year impact (S4)	Treatment	5,168	–0.29	0.96	0.03
		Control	4,991	–0.32	1.02	
Two-year impacts in schools that participated in the two–year study	(4) Two-year novice impact (S5)	Treatment	1,063	–0.18	0.91	0.11
		Control	738	–0.28	1.03	
	(5) Two–year impact (S6)	Treatment	2,631	–0.08	0.93	0.23
		Control	1,145	–0.30	1.02	
	(6) Two-year schoolwide impact (S7)	Treatment	3,486	0.00	0.94	0.14
		Control	3,374	–0.14	1.01	
One–year follow-up impacts in schools that did NOT participate in the two-year study	(7) One-year follow-up novice impact (S8)	Treatment	555	–0.58	0.95	–0.12
		Control	1,161	–0.46	0.98	
	(8) One-year follow-up impact (S9)	Treatment	2,537	–0.51	0.93	–0.18
		Control	3,846	–0.33	1.02	

SD = standard deviation.

Table G.6. Comparison of student baseline grade levels for student impact analyses

Analyses and research questions addressed	Treatment					Control					Is difference > 10%? [effect size]
	Grades: students (percent)					Grades: students (percent)					
	4	5	6	7	8	4	5	6	7	8	
1) One-year novice impact (S1)	124 (5.7)	136 (6.3)	543 (25.1)	735 (34.0)	624 (28.9)	68 (2.6)	285 (10.7)	513 (19.3)	717 (27.0)	1,070 (40.3)	Yes, Grade 8 [-0.31]
(2) One-year impact (S2)	475 (7.7)	740 (12.0)	1,499 (24.4)	1,752 (28.5)	1,684 (27.4)	384 (5.7)	731 (10.9)	1,499 (22.3)	1,805 (26.9)	2,290 (34.1)	No
(3) Overall two-year impact (S4)	265 (5.1)	693 (13.4)	1,282 (24.8)	1,484 (28.7)	1,444 (27.9)	276 (5.5)	569 (11.4)	939 (18.8)	1,744 (34.9)	1,463 (29.3)	No
(4) Two-year novice impact (S5)	0 (0.0)	0 (0.0)	365 (34.3)	308 (29.0)	390 (36.7)	0 (0.0)	57 (7.7)	183 (24.8)	309 (41.9)	189 (25.6)	Yes, Grades 7 & 8 [-0.34 & 0.32]
(5) Two-year impact (S6)	19 (0.7)	364 (13.8)	746 (28.4)	744 (28.3)	758 (28.8)	0 (0.0)	217 (19.0)	366 (32.0)	309 (27.0)	253 (22.1)	No
(6) Two-year schoolwide impact (S7)	42 (1.2)	412 (11.8)	1,131 (32.4)	988 (28.3)	913 (26.2)	56 (1.7)	363 (10.8)	1,112 (33.0)	918 (27.2)	925 (27.4)	No
(7) One-year follow-up novice impact (S8)	52 (9.4)	70 (12.6)	18 (3.2)	214 (38.6)	201 (36.2)	20 (1.7)	167 (14.4)	156 (13.4)	400 (34.5)	418 (36.0)	Yes, Grade 6 [-0.94]
(8) One-year follow-up impact (S9)	246 (9.7)	329 (13.0)	536 (21.1)	740 (29.2)	686 (27.0)	276 (7.2)	352 (9.2)	573 (14.9)	1,435 (37.3)	1,210 (31.5)	No

Note: Percentages do not sum up to 100 percent because of rounding. For dichotomous variables, we assess baseline equivalence using absolute percentage point differences between the intervention and comparison groups with a cutoff of 10 percentage points; this cutoff corresponds to about 0.25 effect size difference when percentages for the two groups are close to 50. Corresponding effect sizes (Cox's indices) are provided in brackets in the last column.

Table G.7. English language arts achievement: alternative specifications for selected student impact models

Model specifications	One-year impact (addressing research question S2)	Overall two-year impact (addressing research question S4)	Two-year impact (addressing research question S6)	One-year follow-up impact (addressing research question S9)
Estimated treatment coefficient (standard error)				
(1) Benchmark (Main model)	0.015 (0.01)	0.084*** (0.02)	0.092*** (0.03)	0.055* (0.03)
(2) Complete case analysis	0.004 (0.01)	0.090*** (0.02)	0.130*** (0.03)	0.038 (0.03)
(3) Student joiners analysis	0.002 (0.01)	0.087*** (0.02)	0.086*** (0.03)	0.056* (0.03)
(4) Using math pre-test instrument for reading pre-test	0.006 (0.02)	0.083*** (0.03)	0.078** (0.03)	0.058 (0.05)
(5) Dropping ELA post-test outliers	0.015 (0.01)	0.084*** (0.02)	0.092*** (0.03)	0.055* (0.03)
(6) Using alternative standardization of outcome	0.020 (0.01)	0.083*** (0.02)	0.083*** (0.03)	0.064* (0.04)
(7) Controlling for school pair dummies	0.005 (0.01)	0.048*** (0.02)	0.055** (0.03)	0.045 (0.04)
(8) Including separate district and grade indicators	0.010 (0.01)	0.071*** (0.02)	0.080*** (0.02)	0.047 (0.03)
(9) Using multiple imputation to impute missing values of independent variables	0.003 (0.01)	0.081*** (0.01)	0.101*** (0.03)	0.043 (0.03)
(10) Using alternative definition of second year of implementation	n.a.	n.a.	0.108*** (0.02)	n.a.

*Statistically significant at the 10 percent level, two-tailed test.
**Statistically significant at the 5 percent level, two-tailed test.
***Statistically significant at the 1 percent level, two-tailed test.
ELA = English language arts, n.a. = not applicable.

B. Supplemental information for student literacy analyses

This section of the appendix discusses the methods used in the impact analysis of student literacy. In the subsections that follow, we first present the predictive validity of the literacy measures, then discuss baseline equivalence of the treatment and control groups, and finally discuss the impact analysis methods and findings.

1. Predictive validity of the literacy measures

To establish predictive validity of literacy measure scores on standardized ELA scores aggregated at the teacher level, we conducted three analyses.

First, we computed bivariate Pearson's correlations between two sets of scores (that is, literacy measure scores and standardized ELA scores aggregated at the teacher level; Table G.8. The correlation coefficient between the two sets of scores is also known as the validity coefficient. The validity coefficient (Pearson's r) can range from –1 to +1. Large coefficients close to 1 in absolute value indicate high predictive validity of the new measure, coefficients close to 0 indicate weak predictive validity and coefficient of 0 indicates no association between the scores. The sample sizes for the analyses ranged from 9 (total score for grades 7–8 and counterclaims) to 12 (all other literacy measures). Missing values were handled by pairwise deletion; if a teacher was missing a score on a literacy measure or a standardized ELA score, that teacher was excluded from the analytic sample for that particular correlation.

Second, to additionally establish the extent to which the scores on literacy measures *jointly* predict scores on the standardized ELA scores, we estimated a set of regression models using literacy measures as predictors and standardized ELA scores as outcomes, additionally controlling for teachers' clustering within schools by using Huber-White robust standard errors. Because of the small sample size (12 teachers clustered in 7 schools), we could not model the joint statistical relationship between all 17 literacy measures and standardized ELA scores. Generally, K+2 observations are required to estimate K parameters in a regression model. Because we only have 7 schools in the sample, we could simultaneously include up to five predictors in a regression model. Therefore, we grouped together literacy measures based on their conceptual meaning in the following order:

1. Organization, voice, sentence variety
2. Vocabulary, spelling, grammar, mechanics
3. Statement of position, reasons, evidence, transition words and links
4. Counterclaims, writing quality total (grades 7–8)
5. Total score (grades 7–8)
6. Writing quality, writing conventions, total score

We additionally computed a seventh measure, the composite score, which is a sum of the following literacy measures: organization, voice, vocabulary, sentence, spelling, grammar, mechanics, statement of position, reasons, evidence, links of transition words, and counterclaims. Internal consistency reliability of the composite scale (Cronbach's alpha) was 0.87.

Table G.8. Pearson's correlation between literacy measure scores and standardized English language arts scores

Literacy measures	Sample size	Pearson's r
Writing conventions domain		
Vocabulary	12	0.19
Mechanics	12	0.44
Spelling	12	0.38
Grammar	12	0.51
Writing conventions total	12	0.50
Writing quality domain		
Voice	12	0.27
Sentence variety	12	-0.01
Organization	12	0.22
Statement of position	9	0.32
Evidence	12	0.49
Reasons	12	0.58
Transition words and links	12	0.06
Counterclaims (grades 7-8)	12	0.13
Writing quality total	12	0.41
Writing quality total (grades 7-8)	9	0.22
Total score	12	0.47
Total score (grades 7-8)	12	0.45

Note: Literacy measures sorted in the order of high to low by the magnitude of Pearson's correlation coefficient within each domain.

Results from the regression analyses are shown in Table G.9. The R^2 statistic shows the total percent of variation in the outcome (standardized ELA scores) jointly explained by literacy measures in each model). Adjusted R^2 statistics provide the proportion of variation in the outcome explained by predictors adjusted for the number of terms in the model. F-statistics provide a global test of significance for all predictors in the model. These statistics show that the predictors explain some of the variance in standardized ELA scores, ranging from R-squared values of 11 to 42 percent, and the F-statistics indicate that predictors have a significant relationship with the standardized ELA scores in three models (model 2, model 3, and model 6).

Note that, because of the sample size, the models have limited power to detect significant effects in the regression models. Statistical power for regression analysis is typically expressed as the probability of a significant finding (that is, a relationship different from zero) when in the

population there is a significant relationship. Statistical power of 0.80 or higher is the standard in applied research and corresponds to at least an 80 percent chance of concluding that there is a real effect. We ran simulation analyses and the statistical power of the student literacy outcome analyses had a joint predictive power between 20 percent (for Model 1 in Table G.9) to 75 percent (for Model 3) to have statistically significant results at the 5 percent level. The small sample sizes in our regression analysis would only be able to detect statistically significant results at the 5 percent level for very large effect sizes (for Model 1 we would need an $R^2 \geq 0.55$ and for Model 3 an $R^2 \geq 0.45$). In other words, having statistically insignificant findings is likely due to the small sample sizes and not necessarily because there was no relationship between the literacy outcomes and the ELA scores

Table G.9. Percent of variation in the standardized English language arts scores jointly explained by literacy measures in three models

Model	Predictors	Sample size	R^2	Adjusted R^2	F statistic
1	Organization, voice, sentence variety	12	0.11	–0.22	0.96
2	Vocabulary, spelling, grammar, mechanics	12	0.28	–0.14	30.36***
3	Statement of position, reasons, evidence, transition words and links	12	0.42	0.08	15.56**
4	Counterclaims, writing quality (grades 7–8)	9	0.39	0.18	5.27
5	Total score (grades 7–8)	9	0.20	0.09	2.85
6	Writing quality, writing conventions, total score	12	0.34	0.09	44.72***
7	Composite score	12	0.13	0.04	5.46

Notes: Negative-adjusted R^2 indicates that the total variation in the outcome explained by predictors is zero or close to zero.

*Statistically significant at the 5 percent level, two-tailed test.
**Statistically significant at the 1 percent level, two-tailed test.
*** Statistically significant at the .01 percent level, two-tailed test.

Third, for each model, we estimated the proportion of variation in standardized ELA scores that is explained by each predictor (literacy measure) separately, when other predictors are held constant in the model (Table G.10). The corresponding statistic is known as partial R^2. We observed small to moderate partial R^2 in the models ranging from 0.00 to 0.37. Like Pearson's r, (partial) R^2 is also a measure of effect size reported in the context of models predicting a variable. The partial R^2 for counterclaims, writing quality (grades 7–8), and total score (grades 7–8) had moderate effect size (> 0.2), whereas the rest of the effects between predictors of standardized ELA scores and outcome were small (< 0.2).

Similar to the discussion of the regression results presented in Table G.9, because of the sample size, these models have limited power to detect significant relationships between the literacy outcomes and ELA test scores and the regression analysis would only be well-powered to detect very large effect size—for example an $R^2 \geq 0.55$—as statistically significant at the 5 percent level.

Table G.10. Percent of variation in the standardized English language arts scores explained by literacy measures in three models

Predictor	Partial R^2	Degrees of freedom	95% Confidence interval of partial R^2	
			Lower boundary	Upper boundary
Model 1				
Organization	0.00	1	.	0.18
Voice	0.05	1	.	0.40
Sentence variety	0.04	1	.	0.39
Model 2				
Vocabulary	0.02	1	.	0.35
Spelling	0.00	1	.	0.25
Grammar	0.07	1	.	0.45
Mechanics	0.00	1	.	0.22
Model 3				
Statement of position	0.00	1	.	0.18
Reasons	0.12	1	.	0.50
Evidence	0.03	1	.	0.39
Model 4				
Counterclaims	0.30	1	.	0.63
Writing quality (grades 7–8)	0.37	1	.	0.67
Model 5				
Total score (grades 7–8)	0.20	1	.	0.56
Model 6				
Writing quality	0.07	1	.	0.43
Writing conventions	0.13	1	.	0.49
Total score	0.05	1	.	0.41
Model 7				
Composite score	0.13	1	.	0.46

Note. The lower boundary of the partial R^2 confidence interval could not be computed because of the small sample size.

2. Baseline differences between treatment and control students at the teacher-level

To test whether there are differences between the treatment and control students' baseline characteristics were aggregated at the teacher level, we estimated a series of t-tests. The t-tests examine differences in students' race and ethnicity, gender, percent of students receiving special education, percent of students who are English-language learners, percent of students eligible for free or reduced-price lunch, and average scores on math and ELA pre-tests. Because of sample size limitations, we could not conduct the t-tests within each grade but for the sample overall and additionally conducted a t-test for the difference between groups in grade level taught. For each t-test, we computed descriptive statistics by group (sample size, means, and standard deviations) and the t-statistic, along with the corresponding *p*-value, effect size (Hedges' g), and the 95 percent confidence interval.

Table G.11 provides results of the t-tests indicating comparisons between the comparison and treatment groups on students' baseline characteristics. We did not observe statistically significant

differences between teachers in treatment and comparison samples at baseline; however, effect sizes greater than the absolute value of 0.25 indicate the presence of differences for several characteristics. In the overall sample, there were differences between treatment and control groups on the proportion of male students, the proportion of students with special education needs, the proportion of students who were English learners, the proportion of students who qualified for free or reduced-price lunch, and students' standardized math pre-test score. Among the grades 7 and 8 subsample, there were differences between the treatment and control groups on all characteristics other than the proportion of students who are black non-Hispanic, the proportion of students who are white non-Hispanic, and the standardized math pre-test scores of students.

Table G.11. Comparison of baseline classroom characteristics between treatment and control group students aggregated at the teacher-level for the literacy analysis

Covariate	Treatment sample	Treatment sample mean	Treatment sample SD	Control sample	Control sample mean	Control sample SD	T–statistic	Effect size	95% C.I. for the effect size	
									Lower boundary	Upper boundary
Overall sample										
Grade	6	6.83	0.75	6	7.00	1.10	0.31	0.16	-0.89	1.21
Average student standardized ELA pre-test score	6	–0.50	0.40	6	–0.54	0.23	–0.23	–0.12	–1.16	0.93
Average student standardized math pre-test score	6	–0.43	0.66	6	–0.68	0.41	–0.78	–0.42	–1.47	0.65
Percent students male	6	0.55	0.07	6	0.51	0.03	–1.22	–0.65	–1.72	0.45
Student race										
Percent black non-Hispanic	6	0.10	0.08	6	0.09	0.14	–0.24	–0.13	–1.17	0.92
Percent white non-Hispanic	6	0.50	0.15	6	0.53	0.19	0.31	0.16	–0.89	1.21
Percent Hispanic	6	0.32	0.18	6	0.29	0.20	–0.30	–0.16	–1.20	0.89
Percent other race	6	0.08	0.06	6	0.09	0.08	0.43	0.23	–0.83	1.27
Percent students eligible for free or reduced-price lunch	4	0.57	0.29	5	0.71	0.19	0.93	0.55	–0.67	1.74
Percent students English language learner	6	0.16	0.17	6	0.09	0.05	–0.87	–0.46	–1.52	0.61
Percent students special education	6	0.16	0.07	6	0.12	0.10	–0.77	–0.41	–1.46	0.66
Grade 7 teachers' sample										
Black non-Hispanic	3	0.05	0.04	3	0.14	0.20	0.71	0.46	–0.88	1.75
Hispanic	3	0.46	0.17	3	0.30	0.29	–0.80	–0.52	–1.82	0.83
White non-Hispanic	3	0.41	0.19	3	0.50	0.19	0.56	0.37	–0.96	1.65
Other race	3	0.08	0.07	3	0.06	0.06	–0.29	–0.19	–1.46	1.11
Male	3	0.56	0.10	3	0.54	0.01	–0.38	–0.25	–1.52	1.06
Special education status	3	0.13	0.09	3	0.12	0.12	–0.10	–0.06	–1.34	1.22
English language learner	3	0.25	0.20	3	0.11	0.06	–1.14	–0.74	–2.07	0.67
Free or reduced-price lunch eligibility	3	0.66	0.27	3	0.79	0.19	0.70	0.45	–0.88	1.74

Covariate	Treatment sample	Treatment sample mean	Treatment sample SD	Control sample	Control sample mean	Control sample SD	T–statistic	Effect size	95% C.I. for the effect size	
									Lower boundary	Upper boundary
Standardized math pre-test score	3	–0.74	0.80	3	–0.72	0.60	0.04	0.03	–1.25	1.30
Standardized ELA pre-test score	3	–0.72	0.42	3	–0.64	0.31	0.27	0.18	–1.12	1.45
Grade 7–8 samples										
Black non-Hispanic	4	0.10	0.10	5	0.10	0.15	0.00	0.00	–1.17	1.17
Hispanic	4	0.40	0.18	5	0.32	0.21	–0.58	–0.35	–1.52	0.85
White non-Hispanic	4	0.44	0.16	5	0.48	0.16	0.31	0.19	–0.99	1.35
Other race	4	0.06	0.07	5	0.10	0.09	0.82	0.49	–0.73	1.66
Male	4	0.55	0.08	5	0.52	0.04	–0.74	–0.44	–1.62	0.76
Special education status	4	0.16	0.09	5	0.09	0.09	–1.10	–0.66	–1.85	0.58
English language learner	4	0.23	0.16	5	0.10	0.05	–1.75	–1.04	–2.29	0.27
Free or reduced-price lunch eligibility	4	0.57	0.29	5	0.71	0.19	0.93	0.55	–0.67	1.74
Standardized math pre-test score	4	–0.69	0.66	5	–0.74	0.43	–0.13	–0.08	–1.25	1.09
Standardized ELA pre-test score	4	–0.69	0.35	5	–0.54	0.26	0.72	0.43	-0.77	1.61

Notes: No findings were statistically significant at the 5 percent level. The measure of the effect size is Hedges' g.

C.I. = confidence interval; ELA = English language arts; SD = standard deviation.

3. Differences between treatment and control students' literacy outcomes at the teacher-level

To test for the differences between treatment and control group students' literacy outcomes aggregated at the teacher-level, we estimated a set of regression models in which literacy task outcomes were regressed on a treatment status indicator and classroom characteristics, controlling for teachers' clustering within schools by using Huber-White robust standard errors. Classroom characteristics included students' race and ethnicity, gender, percent of students receiving special education, percent of students who are English-language learners, percent of students eligible for free or reduced-price lunch, and average scores on math and ELA pre-test. Because of the small sample size (12 teachers clustered in seven schools), we could not accommodate all covariates in the same regression model. Therefore, we estimated a set of regression models that included the following sets of covariates:

1. Treatment status, ELA pre-test scores (Model A)

2. Treatment status, student race and ethnicity (black non-Hispanic, white non-Hispanic, Hispanic, other; Model B)

3. Treatment status, percent students who are English-language learners, percent students receiving special education, percent students eligible for free or reduced-price lunch, math pre-test scores, and percent male students (Model C).

After each regression model, we computed regression-adjusted mean scores for the outcome for treatment and control students at the teacher-level and computed effect sizes (Hedges' g) based on the differences in the regression-adjusted mean scores between the treatment and control groups divided by the observed pooled standard deviation with an adjustment for small sample sizes (Tables G.13–G.15). Note that the regression coefficients are unstandardized and should be interpreted as the magnitude of change in the dependent variable for every one unit change in the corresponding predictor. Because interpretation of unstandardized regression coefficients depends on the metric of dependent and independent variables, one should not formally compare the magnitude of these coefficients. For this reason, rather than interpret the magnitude of effects, we focused on significant findings.

To compensate for the number of inferences being made about instructional practices, we applied the Benjamini-Hochberg multiple comparisons correction to the *p*-values of each pair of impact estimates by subject (Benjamini and Hochberg 1995). Significance levels for adjusted *p*-values are provided in the tables below.

Out of 51 estimated models in Tables G.13–G.15, and after correcting for multiple comparisons, we found no statistically significant differences between treatment and control group students' literacy outcomes at the teacher-level.

Table G.13. Model A: Regression-adjusted differences between treatment and control group students' literacy outcomes at the teacher-level, controlling for baseline ELA pre-test scores

Outcome	Estimated treatment coefficient	Standard error	Effect size	p-value	p-value rank	Adjusted critical value	Significant results?	Analytic sample size	
								Treatment	Control
Writing conventions domain									
Vocabulary	-0.07	0.31	-0.76	0.82	1	0.01	No	6	6
Mechanics	0.05	0.26	0.16	0.85	2	0.01	No	6	6
Spelling	0.01	0.10	-0.01	0.91	3	0.02	No	6	6
Grammar	-0.02	0.24	0.02	0.95	4	0.02	No	6	6
Writing conventions total	-0.02	0.74	-0.01	0.98	5	0.03	No	6	6
Writing quality									
Voice	-0.30	0.10	-2.01	0.02	1	0.01	No	6	6
Sentence Variety	-0.34	0.13	-3.87	0.04	2	0.01	No	6	6
Organization	-0.33	0.35	-1.28	0.39	3	0.02	No	6	6
Statement of position	-0.03	0.09	-0.27	0.76	4	0.02	No	6	6
Evidence	0.18	0.22	0.82	0.44	5	0.03	No	6	6
Reasons	-0.11	0.31	-0.11	0.73	6	0.03	No	6	6
Transition words and links	0.10	0.28	0.59	0.74	7	0.04	No	6	6
Counterclaims (grades 7-8)	-0.14	0.32	-1.41	0.68	8	0.04	No	4	5
Writing quality total	-0.19	1.00	-0.09	0.86	9	0.05	No	6	6
Writing quality total (grades 7-8)	-0.47	1.13	-0.50	0.69	10	0.05	No	6	6
General literacy									
Total score	-0.91	1.73	-0.36	0.62	1	0.03	No	6	6
Total score (grades 7-8)	-1.00	3.61	-0.87	0.80	2	0.05	No	4	5

Note: There were no statistically significant results for any of the literacy task outcomes after multiple comparisons corrections.

C.I. = confidence interval.

Table G.14. Model B: Regression-adjusted differences between treatment and control group students' literacy outcomes at the teacher-level, controlling for students' race and ethnicity

Outcome	Estimated treatment coefficient	Standard error	Effect size	p-value	p-value rank	Adjusted critical value	Significant results?	Analytic sample size	
								Treatment	Control
Writing conventions domain									
Vocabulary	-0.10	0.24	-0.24	0.69	1	0.01	No	6	6
Mechanics	0.10	0.24	0.02	0.69	2	0.02	No	6	6
Spelling	0.13	0.31	0.22	0.69	3	0.03	No	6	6
Grammar	0.05	0.23	-0.02	0.84	4	0.04	No	6	6
Writing conventions total	0.18	0.93	-0.02	0.85	5	0.05	No	6	6
Writing quality									
Voice	-0.32	0.10	-2.04	0.02	1	0.01	No	6	6
Sentence Variety	-0.29	0.19	-2.81	0.18	2	0.01	No	6	6
Organization	-0.38	0.39	-1.35	0.37	3	0.02	No	6	6
Statement of position	0.26	0.33	1.30	0.46	4	0.02	No	6	6
Evidence	-0.76	1.56	-0.76	0.64	5	0.03	No	6	6
Reasons	-0.06	0.14	-0.67	0.67	6	0.03	No	6	6
Transition words and links	-0.19	0.46	-1.22	0.69	7	0.04	No	4	5
Counterclaims (grades 7-8)	0.08	0.19	0.30	0.70	8	0.04	No	6	6
Writing quality total	0.11	0.44	-0.15	0.81	9	0.05	No	6	6
Writing quality total (grades 7-8)	0.00	1.37	-0.13	1.00	10	0.05	No	6	6
General literacy									
Total score	-1.43	5.90	-0.89	0.82	1	0.03	No	4	5
Total score (grades 7-8)	-0.51	2.35	-0.60	0.84	2	0.05	No	6	6

Note: There were no statistically significant results for any of the literacy task outcomes after multiple comparisons corrections.

C.I. = confidence interval.

Table G.15. Model C: Regression-adjusted differences between treatment and control group students' literacy outcomes at the teacher-level, controlling for students' standardized math pre-test scores, English-language learner status, special education status, free or reduced-price lunch status, and gender

Outcome	Estimated treatment coefficient	Standard error	Effect size	p-value	p-value rank	Adjusted critical value	Significant results?	Analytic sample size	
								Treatment	Control
Writing conventions domain									
Vocabulary	-0.47	0.18	-0.14	0.04	1.00	0.01	No	6.00	6.00
Mechanics	-0.05	0.38	0.23	0.90	4.00	0.02	No	6.00	6.00
Spelling	0.04	0.46	-0.01	0.93	5.00	**0.03**	No	6.00	6.00
Grammar	-0.14	0.39	0.02	0.72	3.00	**0.02**	No	6.00	6.00
Writing conventions total	-0.62	1.31	-0.01	0.65	2.00	0.01	No	6.00	6.00
Writing quality									
Voice	-0.58	0.16	-0.84	0.01	1.00	0.01	No	6.00	6.00
Sentence Variety	-0.70	0.22	-0.75	0.02	2.00	0.01	No	6.00	6.00
Organization	-1.09	0.37	-0.49	0.02	3.00	0.02	No	6.00	6.00
Statement of position	-0.45	0.23	-0.37	0.12	4.00	0.02	No	4.00	5.00
Evidence	-2.89	1.64	-0.23	0.13	5.00	0.03	No	6.00	6.00
Reasons	-0.36	0.24	0.22	0.18	6.00	0.03	No	6.00	6.00
Transition words and links	-2.07	1.62	-0.05	0.25	7.00	0.04	No	6.00	6.00
Counterclaims (grades 7-8)	-0.13	0.15	-0.31	0.40	8.00	0.04	No	6.00	6.00
Writing quality total	-0.37	0.68	-0.11	0.60	9.00	0.05	No	6.00	6.00
Writing quality total (grades 7-8)	-0.11	0.43	0.59	0.80	10.00	0.05	No	6.00	6.00
General literacy									
Total score	-6.00	3.40	-0.44	0.15	1.00	0.01	No	4.00	5.00
Total score (grades 7-8)	-4.05	2.99	-0.25	0.22	2.00	0.01	No	6.00	6.00

Note: There were no statistically significant results for any of the literacy task outcomes after multiple comparisons corrections.

C.I. = confidence interval.

Mathematica

Princeton, NJ • Ann Arbor, MI • Cambridge, MA
Chicago, IL • Oakland, CA • Seattle, WA
Tucson, AZ • Woodlawn, MD • Washington, DC

EDI Global, a Mathematica Company

Bukoba, Tanzania • High Wycombe, United Kingdom

mathematica-mpr.com