

Project Narrative

Contents

	Page
Introduction.....	1
A. Significance.....	3
A.1. Severity of the Problem	3
A.2. National Significance.....	3
A.3. Exceptional Approach to Absolute Priority 2.....	5
B. Strategy to Scale.....	9
B.1. Demand for Professional Learning and Support for Instruction.....	9
B.2. A Specific Strategy to Scale That Addresses Barriers to Quality.....	12
B.3. Feasibility of Successful Replication in a Variety of Settings and Populations	15
C. Quality of the Project Design and Management Plan	16
C.1. Clearly Specified and Measurable Goals, Objectives, and Outcomes	16
C.2. A Management Plan Defining Responsibilities, Timelines, and Milestones.....	18
C.3. Procedures Ensuring Feedback and Continuous Improvement	22
C.4. Ongoing Work Beyond the End of the Grant.....	23
D. Quality of the Project Evaluation.....	24
D.1. Evaluation Methods Designed to Meet WWC Evidence Standards Without Reservations.....	26
D.2. Generation of Guidance About Effective Strategies Suitable for Replication	26
D.3. Valid and Reliable Performance Data on Relevant Outcomes	28
D.4. Clear Articulation of Components, Mediators, and Outcomes and Measurable Threshold	30
References.....	31

Introduction

The need for coherent, effective, and sustainable approaches to improving instruction for high-need students cannot be overstated. Results from the 2017 National Assessment of Educational Progress (NAEP), for example, show that students from low-income families and students of color continue to achieve at far lower rates than their fellow students and that both reading and mathematics achievement among our lowest performing fourth graders has been declining (NAEP, 2018). The causes for these gaps are many, but closing them will require bringing to scale programs that effectively help teachers who serve underperforming students implement better instruction every day. The *Professional Learning with Impact* (PLI) program is one such program.

The American Institutes for Research (AIR), The Danielson Group, and Learning Forward propose a project to improve, test, and scale the PLI program, which integrates The Danielson Group's proven one-on-one coaching model and Learning Forward's distinctive approach to teacher learning teams (TLTs). A fourth partner, Educopia, will provide the online tools needed to implement PLI. Six partner districts are eager to participate, as PLI provides multiple supports aligned to the rubric all of them use for teacher evaluation—Charlotte Danielson's Framework for Teaching (FFT)—and promises to build teacher capacity to implement their current curricula effectively. The districts will focus on school-based teams of teachers in Grade 4, who share the same curriculum as well as school and grade-level goals, to set the stage for future scaling across all grade levels.

To implement the program, The Danielson Group and Learning Forward will train, monitor, and support selected instructional leaders in each district to be the PLI coaches, each serving the Grade 4 teachers at two or three schools. Each teacher will receive a combination of summer

workshops, one-on-one coaching, and facilitated learning team meetings, totaling 90 hours across 2 years.

To facilitate iterative evaluation and refinement of PLI, AIR’s partners will implement the 2-year PLI program in three successive cohorts, as shown in Exhibit 1, focusing on elementary schools with high percentages of high-need students. Each cohort will include diverse school settings and district contexts. In their work with each cohort, AIR’s partners will use feedback routines, internal data, and independent evaluation data from AIR to improve the PLI program continuously and revise their manuals and other supporting materials for future implementations.

Exhibit 1. Number of Schools to Receive the PLI Program, by Cohort and Year

School Cohort	2018–19 (Y1)	2019–20 (Y2)	2020–21 (Y3)	2021–22 (Y4)	2022–23 (Y5)
Cohort 1 (N=6)		√	√		
Cohort 2 (N=12)			√	√	
Cohort 3 (N=24)				√	√

In addition, AIR will conduct an experiment to determine PLI’s impact on key teacher and student outcomes. For Cohorts 2 and 3 specifically, AIR and the districts will identify twice as many schools as needed. Then, within each district, AIR will randomly assign half of the identified schools to treatment and half to control and collect data from both groups (e.g., Cohort 2 will have 12 treatment schools, as shown in Exhibit 1, and 12 control schools).

The project is fully aligned with the EIR Mid-Phase program, especially as it addresses *Absolute Priority 2—Field Initiated Innovations-General* by implementing, improving, and scaling the innovative, evidence-based PLI program in schools with high concentrations of high-need students.

A. Significance

A.1. Severity of the Problem

The problem that this project is designed to address—persistent gaps in student achievement based on family background—is severe. For example, the 2017 NAEP showed that nearly one in two Black fourth graders (49%) were reading below the basic level (NAEP, 2017b)—more than twice the proportion of White students who scored below basic (22%). In mathematics, 31% of fourth graders from low-income households scored below basic, compared to just 9% of their higher income counterparts. These gaps in achievement do not go away by Grade 8, or in high school, and they cut across other segments of the U.S. student population—such as city and suburban students, Latino/a and white students, and students with disabilities and those without (National Center for Education Statistics, 2017; NAEP, 2017a).

In addition to being severe, these gaps are also stubborn. The achievement gaps in Grade 4 reading and math between White and Black students, for example, have remained essentially static for the last decade (NAEP, 2017a). The consequences are long-term and far-reaching: low levels of literacy and numeracy systematically limit high-need students in educational attainment and long-term health and economic well-being (Baker, Wolf, & Feinglass, 2007) and even limit the economic productivity of the nation as a whole (Auguste, Hancock, & Laboissiere, 2009).

A.2. National Significance

The proposed project is significant because it has the potential to reduce achievement gaps and improve achievement for high-need students. It focuses on schools with high percentages of students who are Black or Hispanic (on average 85%; see Appendix G.3) and from low-income families (on average 88%) and addresses two areas of critical national need: teacher quality and support for teachers to design and deliver effective instruction.

Teacher quality is seen by researchers and policymakers as a potential lever for improving students' engagement, achievement, and later life outcomes (see e.g., Chamberlain, 2013; Chetty, Friedman, & Rockoff, 2014; Gershenson, 2016; Rivkin, Hanushek, & Kain, 2005). Boosting teacher quality may be especially important for students from low-income families and students of color, who face gaps in access to effective instruction (Cowan, Goldhaber, & Theobald, 2017; Goldhaber, Lavery, & Theobald, 2015; Goldhaber, Quince, & Theobald, 2016; Isenberg et al., 2016; Sass, Hannaway, Xu, Figlio, & Feng, 2012). In fact, the *Every Student Succeeds Act* calls on states and districts to ensure that such students are not taught disproportionately by ineffective teachers.

However, effective approaches to improving teacher quality are hard to find. For example, researchers have found that some teacher professional development programs have an impact on student achievement, but many do not. For program reviews, see Blazar, Kraft, and Hogan (2017); Desimone and Garet (2015); Garet, Heppen, Walters, Smith, and Yang (2016); Gersten, Taylor, Keys, Rolfhus, and Newman-Gonchar (2014); and Kennedy (2016). The scarcity of effective approaches to professional learning is puzzling given the crucial role that improvements in employee knowledge and skill play in other professions.

The second area of critical national need—support for teachers to design and deliver effective instruction—has emerged as an urgent challenge as states and districts adopt new curricula based on updated college- and career-readiness standards (Hirsh, 2018; Olson, 2018). Evidence shows that some high-quality curricula can improve student achievement (e.g., Agodini, Harris, Thomas, Murphy, & Gallagher, 2010; Borman, Dowling, & Shneck, 2008), but good curriculum is not effective on its own (Chingos & Whitehurst, 2012). To translate it into instruction, teachers do not simply follow the materials that accompany the curriculum; they

draw on professional knowledge and understanding, school- and grade-level goals, and knowledge of their students (Cohen & Ball, 1999; Shulman, 1987), often creating their own materials or adapting materials developed by colleagues (Kane, Owens, Marinell, Thal, & Staiger, 2016; Opfer, Kaufman, & Thompson, 2016). There is a critical need to support teachers in designing their instruction to ensure that they succeed and increase student learning (Dysarz, 2018).

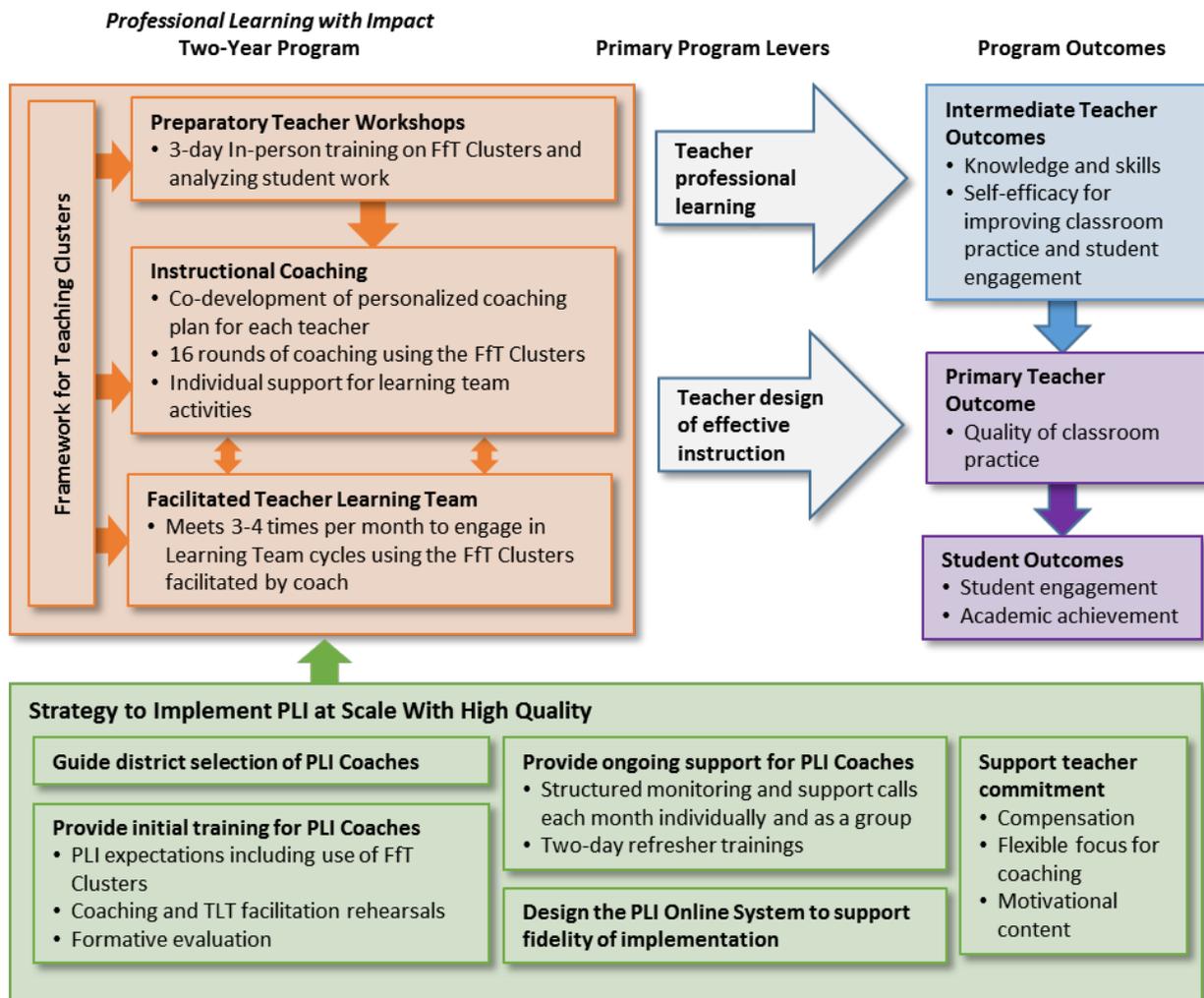
PLI is designed to address both areas of critical national need. Through individualized instructional coaching, PLI stimulates professional learning and makes teachers more effective. Through TLTs, it engages teachers in sustained efforts to design and deliver effective instruction, with facilitated support from team members in the same school and grade level, leading to improved classroom practice. PLI integrates coaching and TLTs, grounding them in the FfT, fulfilling a need for aligned, ongoing, and mutually supportive practices for professional learning and improved instruction (Hirsh & Crow, 2017; Weiner & Pimental, 2017).

A.3. Exceptional Approach to Absolute Priority 2

This project represents an exceptional approach to Absolute Priority 2 because of its focus on the field-initiated PLI program, which is highly **innovative** and built on a theory of action (see Exhibit 2) that is well-grounded in rich **evidence**, as described in this overview of the three main components of PLI (depicted in the orange box in Exhibit 2): preparatory teacher workshops, instructional coaching, and facilitated TLTs.

Preparatory Teacher Workshops. To prepare for their role in PLI, treatment teachers will participate in a 3-day in-person workshop at the beginning of the program, focusing on using the newest version of the FfT called the FfT Clusters to analyze instruction and student work (see Appendix G.1.c for samples of the workshop materials). The FfT Clusters incorporates two innovations. First, based on feedback that the original, 22-dimension FfT felt overwhelming to

Exhibit 2. PLI Theory of Change



some teachers, this version taps the same underlying dimensions but consolidates them into six, readily-interpreted clusters. In addition, based on expectations embodied in college- and career-readiness standards, the FFT Clusters include content-specific language about instruction in English language arts (ELA) and mathematics, to make it easier for teachers and coaches to apply the FFT Clusters in these instructional contexts. For example, the mathematics version refers to the use of representations and modeling of mathematical language, in addition to more general features of instruction. (See Appendix G.1.a for the full FFT Clusters – Math Version.)

Using the latest version of the FFT, PLI is grounded in solid **evidence** on FFT’s positive effects on achievement. Two multisite randomized experiments described on the Evidence Form

for this application (Garet et al., 2017; Steinberg & Sartain, 2015) and one quasi-experimental study (Taylor & Tyler, 2012) found that providing teachers with feedback using the FFT led to improved student achievement. Researchers also have found positive relationships between teachers' FFT scores and student achievement gains (e.g., Kane & Staiger, 2012).

Instructional Coaching. In the PLI program, selected and trained district-based PLI coaches develop a coaching plan based on each teacher's identified needs and engage in 16 rounds of video-enabled coaching, roughly one per month during the 2 school years of the program. To leverage the benefits of video technology, each standard round of coaching follows an **innovative** process similar to that used in the coaching program MyTeachingPartner. The teacher video records a 15–20-minute sample of his or her instruction intended to provide a basis for discussing one or two teaching skills or elements of practice in the FFT Clusters; for some elements, the teacher also uploads lesson plans and assignments. Using the PLI web-based platform, the PLI coach chooses two to three short video clips of important illustrative moments and writes narrative observations or questions about the clips using language from the FFT Clusters. After the teacher watches the clips and reads the coach's response, the coach and teacher engage in a learning-focused conversation to clarify and solidify the teacher's learning and guide immediate improvements to practice. (For screenshots of the PLI Online System to support the coaching and sample materials from the PLI coach training, see Appendix G.1.b and G.1.e.)

The **evidence** from impact studies of the coaching program MyTeachingPartner demonstrates that coaching using video clips and a rubric to guide reflection can improve classroom practice and student achievement (Allen, Pianta, Gregory, Mikami, & Lun, 2011; Allen, Hafen, Gregory, Mikami, & Pianta, 2015). More broadly, a meta-analysis of 44 empirical studies by Kraft,

Blazar, and Hogan (2018) finds a positive average effect of instructional coaching on instruction and achievement (with pooled effect sizes of .58 on instruction and .15 on achievement).

Teacher Learning Teams. TLTs use **innovative** strategies to promote both teacher learning and collaborative instructional design grounded in knowledge of the curriculum and a teacher’s students (Hirsh, 2018). For the TLT component of PLI, teachers meet weekly in facilitated grade-level PLCs, led by their PLI coach. (The role of the PLI coach spans instructional coaching and TLT facilitation, to ensure integrated support.) Similar to many collaborative inquiry approaches, the TLTs engage in Learning Team cycles that each span several weeks. To structure and support these cycles, PLI specifies five stages of inquiry and design (Hirsh & Crow, 2017), grounded in the FfT Clusters as shown in Exhibit 3.

Exhibit 3. Five Stages of the Learning Team Cycle

- | |
|---|
| 1. Teams examine student and educator learning challenges by analyzing student work, data, curricular materials, learning expectations, and other information. |
| 2. Teams examine their findings from Stage 1 through the lens of the FfT Clusters and identify shared goals for student and educator learning. |
| 3. Teams collaborate to explore or innovate solutions to those challenges using the FfT Clusters and other resources provided by the PLI coach or one another, sometimes simulating those solutions with their colleagues to get additional FfT-based feedback. |
| 4. Teachers implement their new learning and instructional designs in classrooms and collect information on students’ responses. |
| 5. Teams examine students’ responses and teachers’ observations and adjust their classroom practice as a result, documenting their learning as they go. |

These five stages put student learning and the curriculum at the center of the teacher’s efforts and guide professional learning, with the FfT Clusters helping teachers co-develop and refine instructional practices. In addition to the social support for this work (Hirsh & Crow, 2017), a teacher can ask the PLI coach to provide one-on-one support, helping the teacher try a new practice, reflect, and consider additional strategies.

The **evidence** for this kind of rigorous, inquiry-based teacher collaboration shows that it, too, holds promise for improvement in teaching and student outcomes (Ermeling, 2009; Lomos, Hofkin & Bosker, 2015; Saunders, Goldenberg, & Gallimore 2009). In general, reviews of this research have found wide variation in how teacher teams are conceptualized and implemented but many benefits to collaboration, such as reducing teacher isolation, increasing teacher efficacy and morale, and increasing instructional innovation (DeLuca et al., 2015; Vangrieken et al., 2015).

In sum, the proposed project will improve and test an **exceptionally innovative and evidence-based** teacher professional development (PD program, PLI, that combines instructional coaching and curriculum-focused teacher learning teams using a common instructional framework. As shown in the theory of change in Exhibit 2, the PLI program’s primary levers are (1) teacher professional learning and (2) teacher design of effective instruction. These are designed to improve classroom practices and student achievement in the partner districts. The project also will show how PLI can be brought to scale.

B. Strategy to Scale

B.1. Demand for Professional Learning and Support for Instruction

The project will reach the intended level of scale—42 schools by Year 4—and spread further because of strong demand for effective professional learning systems and widespread popularity for coaching and PLCs as modes of PD. Districts invest significant resources in PD, including billions of dollars from federal programs (U.S. Department of Education, 2014). A recent study of three large districts found that mandatory PD for teachers required from one to two weeks each year, and spending on teacher PD was at least 5% of annual spending (Hasiotis, Jacobs, & McGovern, 2015). To build local instructional capacity, more and more districts spend these resources on instructional coaching, PLCs, or both. According to the latest available survey

tabulations, 66% of U.S. public schools had staff in coaching assignments (NCES, n.d.), and 81% of all teachers reported having participated in regularly scheduled collaboration on instructional issues (Rotermund, DeRoche, & Ottem, 2017).

What districts demand beyond these general modes of PD—and what makes the PLI program distinctively appealing—is effectiveness, scalability, and coherence.

- *Effectiveness.* Grounded in the FfT, feedback from which is proven to have an impact on achievement (see Section A.3), the PLI program meets a need for effective PD. The Every Student Succeeds Act adds pressure on districts to show that they use proven programs. Yet awareness is spreading that few teacher PD programs show significant impact on achievement when tested in rigorous studies (e.g., Garet, Heppen, Walters, Smith, & Yang, 2016; see Section A.2). In addition, although a meta-analysis of instructional coaching studies (Kraft et al, 2014) concluded that coaching programs had a significant average impact across all studies reviewed, there are only a small number of coaching programs that have demonstrated significant impact in an individual study.
- *Scalability.* PLI is well-suited for scaling, including clear protocols and expectations for the coaching and TLT facilitation as well as strong supports for selection, training, and ongoing monitoring of the PLI coaches (for details, see Section B.2). Coaching and TLT facilitation require special skills, and many districts struggle to provide coaches and PLC facilitators with clear expectations and ongoing support (Neufeld & Donaldson, 2012; Woulfin & Rigby, 2017). Commenting on the challenge of scaling instructional coaching, Kraft et al. (2018) noted that the studies that provided coaching to 100 or fewer teachers yielded double the impact on achievement of those with more than 100 teachers (p. 29).

Thus, there is a clear need for effective and scalable coaching programs—a need that PLI is intended to fulfill.

- *Coherence.* Coherence is widely seen as a critical feature of effective PD (e.g., Darling-Hammond, Wei, Andree, Richardson, & Orphanos, 2009; Desimone et al., 2002) and is also a distinctive feature of PLI. PLI provides a coherent package of individual and group support for teachers because all program components are grounded in a common framework (the FfT) (see Section A.3.) and focus teachers on the district’s curriculum. Moreover, because the most common instrument for districtwide teacher evaluation in the United States is the FfT, or an adaptation of it, many districts see PLI as being aligned to their district and school priorities.

These features of PLI have attracted six districts from six states to volunteer to participate in the project. All six districts (1) use the FfT as part of their teacher evaluation and support system; (2) have tried PLCs but lack a promising, well-specified model; (3) want to support teachers’ efforts to implement the curriculum; and (4) consider high-need students a priority. The districts include an ample supply of elementary schools with high percentages of high-need students (659 schools in which 75% or more students eligible for free- or reduced-price lunch; see Appendix D for letters of support and Appendix G.3 for demographic characteristics for this subset of schools). In addition, the Connecticut State Department of Education will facilitate recruitment of a group of small rural districts (see Appendix D).

This surplus of schools with high-need students gives the project flexibility to fill each cohort strategically, balancing across district and school settings in terms of size and other factors (e.g., geography, student race and ethnicity) to generate valuable guidance for later replications. If additional districts are needed during the project for any reason, we will use multiple methods to

recruit them. For example, AIR will ask the partner districts to introduce AIR to their counterparts elsewhere, allowing officials in prospective districts to hear first-hand from trusted contacts about what participation in the project would mean. The Danielson Group and Learning Forward will tap into their national networks to identify potential partner districts to introduce to AIR and conduct sessions at their annual conferences for districts interested in learning more about trying PLI.

B.2. A Specific Strategy to Implement PLI at Scale With High Quality

In the literature on implementation of coaching and PLCs, there are some key barriers to scaling with high quality. To overcome these barriers, the project leaders at Danielson Group and Learning Forward—known as the lead coach (LC) and lead facilitator (LF), respectively—will execute several specific strategies, as discussed in this section.

Guide District Selection of Coaches. A common barrier to successful coaching and PLC programs is poor selection of staff to serve as coaches and facilitators (New Teacher Center, 2016). Therefore, a key feature of the PLI scaling strategy is the guided selection of skilled PLI coaches. The LC will provide the school district partners with a list of guidelines for desired qualities for PLI coaches, which include, for example, experience teaching in an elementary school; evidence of supportive relationships with teachers; and strong interpersonal skills to facilitate collaborative, teacher-driven conversations. The school district partners will select additional personnel to serve as back-up coaches in case of staff mobility or turnover, another common barrier to effective implementation and scale-up of coaching programs.

Provide Initial Training for PLI Coaches. Another common barrier to high-quality coaching and PLC implementation is lack of clear expectations (see Section B.1) and sufficient preparation for the coaches. To address this barrier, all PLI coaches will participate in six-days of intensive workshops, where PLI leads will specify and unpack the expectations for PLI coaches.

The workshops also will incorporate rehearsals and formative evaluation to ensure that all the PLI coaches are able to meet program expectations (see Appendix G.1.b for sample materials developed by the Danielson Group and Learning Forward). In addition, the coaches will attend a 2-day refresher workshop at the start of the second intervention year.

Provide Ongoing Support for PLI Coaches. To further ensure that PLI coaches fulfill the expectations and support teachers successfully, the LC and LF will provide ongoing support for the coaches through regular, structured check-ins. These will include monthly “check-in” meetings with each coach individually and separate monthly group meetings (Exhibit 4), which will become bimonthly after one year of implementing PLI coaching with fidelity. The individual meetings will allow the LC and LF to monitor and support each coach. Before a meeting, the LC and LF will use a rubric (see Appendix G.1.d) to review materials captured in the online system for one recent coaching round and one recent TLT meeting. The LC and LF will then share the completed rubrics with the coach at the meeting and offer individualized support focused on identified challenges. For group check-ins, the LC and LF will co-facilitate a webinar in which PLI coaches will take turns sharing their questions and experiences. Support will come from the LC and LF as well as the other PLI coaches in the webinar, who will share implementation lessons from other schools and districts. The LC and LF also may use these check-ins to gain feedback on project implementation.

Exhibit 4. Ongoing Support Check-Ins During a PLI Coach’s First Year of Coaching

Ongoing Support Activity	Mode of Ongoing Support	Duration (hours)
Group Support Check-ins	Group webinars offered to all PLI coaches	1.5 hours/month
Individual Support Check-ins	Phone calls with individual PLI coaches	1 hour/month

To illustrate a barrier that a PLI coach may need support to overcome, coaches often struggle to maintain a positive, trusting relationship with teachers while providing feedback that is critical (Neufeld & Donaldson, 2012). Similarly, as a TLT facilitator, a PLI coach also may struggle

with getting teachers to strike this balance when supporting each other. The PLI coaches will rehearse such challenging conversations in their preparatory trainings, but if several struggle, the LC and LF will use a group support check-in to provide general support and encourage collaborative problem solving. If a PLI coach is having especially difficult challenges with relationships with teachers, the LC will offer one-on-one support as needed.

Ensure Teacher Commitment to PLI. We plan specific steps to ensure teacher commitment to engaging in PLI: (1) reimburse teachers for their time; (2) obtain advanced approval for continuing education credits for participation, (3) incorporate motivational content in the Preparatory Teacher Workshops (see Section A.3), and (4) give teachers flexibility about the focus of the coaching. Because the coaching and TLTs are *directly connected* to teacher’s day-to-day practice—as well as their formal evaluation system—teachers will find the program relevant to their work, rather than distracting.

Provide Summer Workshops Directly to Teachers. The PLI summer trainings for teachers, based on existing materials from The Danielson Group and Learning Forward, provide the foundational tools needed to benefit from PLI. To reach the intended scale with quality, the LC and LF will provide these workshops directly to teachers, with separate events in each district. Thus, teachers will learn first-hand about the FfT Clusters and the PLI protocols for analyzing student work and determining promising instructional strategies. Coaches will attend these summer workshops for teachers (after having participated in their own preparatory workshops) to reinforce their own learning and help the LC and LF articulate the relevance of the PLI program for teachers in the local district context.

Design the PLI Online System to Support Fidelity of Implementation. The PLI online system will support fidelity by structuring the exchange of classroom videos, video clips, and messages between teacher and coach described in Section A.3. (For screenshots, see Appendix G.1.e.)

The partners will use these supports to scale PLI with quality. The budget includes resources to pay a share of cost of PLI coach labor, with the school district partners paying the remainder. We expect the project's 168 treatment teachers to be served by a combination of full- and part-time district-based coaches. One full-time PLI coach can provide instructional coaching and TLT facilitation to three schools (approximately 12 Grade 4 teachers).

B.3. Feasibility of Successful Replication in a Variety of Settings and Populations

If the project's evaluation demonstrates an impact on student achievement, it will be feasible to use the scaling strategies to deliver PLI in a variety of settings and populations, for several reasons. First, this project will result in a well-developed set of materials—including the online system—to support the roles of the LC and LF, PLI coaches, and teachers (see Section B.2).

These materials will be iteratively refined across three cohorts that include a variety of settings and populations, to incorporate useful lessons to inform future replications of the PLI program.

Second, to further increase the feasibility of replication, Learning Forward will convene the participating districts five times: twice as a community of practice to discuss feedback on the PLI program and its implementation and three times as part of a broader community of practice. This broader community will include a network of Learning Forward districts that use the FfT and have been meeting to share lessons about creating coherent systems of support for teachers. The meetings will engage users of PLI, such as teachers, coaches, and central office staff, to help them support each other with successful implementation. In addition, they will support districts that are considering adopting PLI and want to hear current users' perspectives.

C. Quality of the Project Design and Management Plan

C.1. Clearly Specified and Measurable Goals, Objectives, and Outcomes

The overall goal of the project is to further develop, test, and implement at scale the PLI program and to determine its impact. Exhibit 5 lists the specific objectives for the project, strategies for achieving the objectives, expected outcomes, and measures of the outcomes. Unless otherwise specified, the strategies are to be used for all three cohorts.

Exhibit 5. Objectives, Strategies, Outcomes, and Measures

Strategies	Outcomes	Measures
Objective 1. Deliver supports for <i>Strategy to Implement PLI at Scale With High Quality</i> while continuously using feedback and fidelity data for project improvement.		
Strategy 1.1. Guide recruitment of district-based PLI coaches.	District staff provided with PLI coach hiring criteria and trained for their role in coach recruitment.	Measure 1.1. All interviewers attend 95% or more of the training. Districts submit plans for coach recruitment.
Strategy 1.2. Recruit local staff to serve as PLI coaches (responsible for instructional coaching and TLT facilitation).	PLI coaches with sufficient qualifications and skills who understand the expectations for their role in PLI and are committed to implementing PLI.	Measure 1.2. Based on coach applicant screening and interview records, 100% of PLI coaches hold the minimum qualifications or better and complete an interview.
Strategy 1.3. Conduct preparatory workshops for PLI coaches.	PLI coaches prepared to provide individualized instructional coaching, facilitate TLTs, and use the PLI online system.	Measure 1.3. All PLI coaches attend 95% of the 6 days of training; coach workshop exit surveys identify no critical gaps.
Strategy 1.4. Provide ongoing monitoring and support for PLI coaches through Group Support and Individual Support meetings and one-on-one assistance as needed.	PLI coaches are monitored and receive peer group and individualized support in their role.	Measure 1.4. Coach support meeting attendance records indicate each coach attends at least 67% of meetings; additional measures based on coach, LC, and LF interviews.
Strategy 1.5. Refine materials and procedures for each strategy under Objective 1.	Improved materials: PLI coach manual; PLI Teacher manual.	Measure 1.5. Biannual memo summarizing revisions made to materials and procedures for Objective 1 contains all required elements.

Strategies	Outcomes	Measures
Objective 2. Implement PLI while continuously using feedback and fidelity data for project improvement.		
Strategy 2.1. Conduct district-based preparatory workshops for treatment teachers.	Treatment teachers understand the FfT Clusters and how to use protocols for analyzing student work and commit to participating in PLI. Treatment teachers are prepared to engage in instructional coaching and TLT activities.	Measure 2.1. Based on teacher workshop attendance records, 100% of teachers attend 85% of the 3-day workshop; teacher exit surveys identify no critical gaps.
Strategy 2.2. Provide eight rounds of instructional coaching per teacher each year for 2 years.	PLI coaches complete intended number of rounds with each treatment teacher with fidelity; treatment teachers report more coaching than control; impacts occur on quality of classroom practice, student engagement, and student achievement.	Measure 2.2. Based on online coaching activity trackers, each teacher completes at least 75% of the planned 16 coaching cycles. Each coach demonstrates proficiency on 80% of items from the fidelity of implementation (FOI) rubric for coaching.
Strategy 2.3. Facilitate weekly TLT meetings including up to eight learning team cycles.	PLI coaches facilitate weekly TLT meetings attended by all treatment teachers; treatment teachers report better-focused team meetings than control; impacts occur on quality of classroom practice, student engagement, and student achievement.	Measure 2.3. Based on online TLT activity trackers and posted materials, each PLI coach facilitates TLT meetings three times per month during school year, and each teacher attends 75% or more of the meetings. Each coach demonstrates proficiency on 80% of items from the FOI rubric for TLTs.
Strategy 2.4. Refine materials and procedures for each strategy under Objective 2.	Improved materials: PLI coach manual; PLI teacher manual; PLI online system.	Measure 2.4. Biannual memo summarizing revisions made to materials and procedures for Objective 2 contains all required elements.
Objective 3. Conduct an implementation evaluation (all cohorts) and RCT to test the impact of PLI on the quality of instruction and student outcomes (Cohorts 2 and 3 only).		
Strategy 3.1. Identify participating schools in partner districts and recruit principals and teams of Grade 4 teachers in those schools.	Principals at schools identified by partner districts support the project and agree to be randomly assigned.	Measure 3.1. District and school signatures on project memo of understanding (MOU) for 6, 24, and 48 schools for Cohorts 1, 2, and 3, respectively.
Strategy 3.2. Randomly assign schools to treatment and control conditions.	Samples of treatment and control schools with baseline equivalence in key student and teacher characteristics.	Measure 3.2. Random assignment and baseline equivalence memo report on all baseline measures

Strategies	Outcomes	Measures
Strategy 3.3. Measure and analyze fidelity of implementation.	Data on fidelity of implementation collected and analyzed.	Measure 3.3. Data collection update indicates response rate of 90% or better; fidelity memo for each strategy that supports implementation (i.e., Strategies 1.1-1.4 and 2.1-2.3).
Strategy 3.4. Measure and analyze treatment-control contrast in teachers' PD experiences.	Data on frequency and content of teachers' PD experiences collected and analyzed.	Measure 3.4. Data collection update indicates response rate of 90% for teachers.
Strategy 3.5. Assess the impact of PLI on classroom practice, student engagement, and student achievement.	Data on outcome measures collected and analyzed.	Measure 3.5. Data collection update indicates response rate of 90% for teachers and 80% for students; impact memo that meets WWC standards without reservation.
Objective 4. Develop sample and infrastructure for continued scaling.		
Strategy 4.1. Identify more districts if needed.	Commitment of districts to participate in project.	Measure 4.1. Additional signed MOUs (see Measure 3.1).
Strategy 4.2. Convene meetings to support and sustain work of participating districts and other districts with interest in PLI (Years 1–5).	PLI coaches and district leaders share experiences that help other users and inform those considering implementing PLI.	Measure 4.2. Meeting attendance records indicate all partner districts attend; 4 additional districts attend; memo summarizing input from partners and lessons.

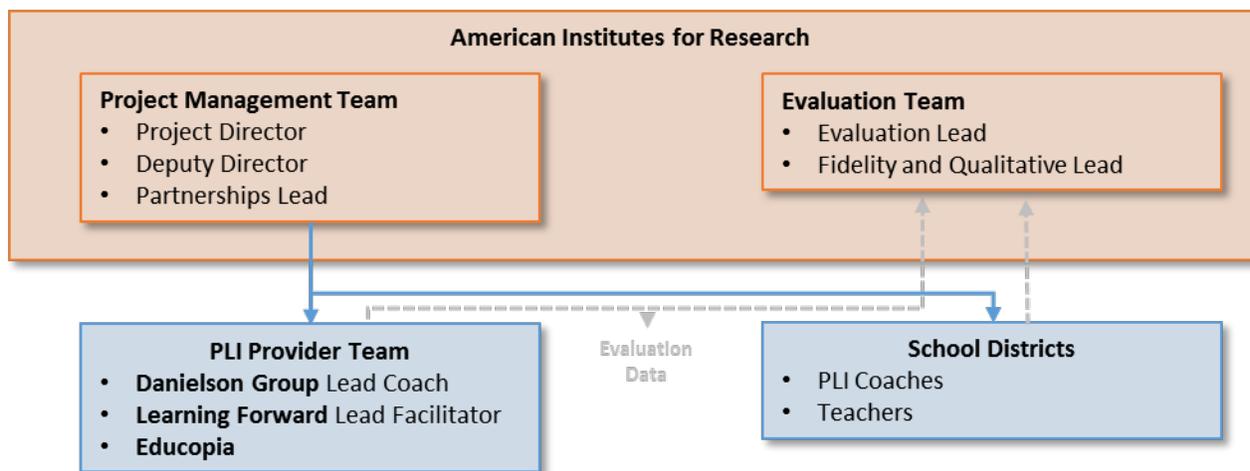
C.2. A Management Plan Defining Responsibilities, Timelines, and Milestones

The management plan establishes the reporting relationships for the partner organizations (see Exhibit 6). The plan is more than adequate in part because each partner organization is highly qualified for a clear and specific role involving execution of the strategies (see Section C.1) at each milestone on the project's 5-year timeline (see Exhibit 8).

AIR is the lead organization for the project, responsible to the U.S. Department of Education for grant performance. AIR's role is to (1) oversee the subgrants to the service providers (Danielson Group, Learning Forward, and Educopia) and to the school district partners, ensuring coordination across the partners to achieve the project objectives; (2) recruit eligible schools from the school district partners; and (3) conduct the independent evaluation. To ensure the

independence of the evaluation, the AIR evaluation team will be separate from the AIR project management team, as shown in Exhibit 6, and will have no role in the development or the implementation of the PLI intervention except to share implementation analyses as feedback. This structure ensures the independence of key evaluation activities including random assignment, outcome data collection, analysis, and reporting, and is consistent with OII guidance (U.S. Department of Education, Office of Innovation and Improvement, n.d.; Abt Associates, 2015). In addition, AIR trains staff to report concerns about independence and tracks labor charges by task to ensure that team members follow the intended division of labor between the project management tasks and the evaluation tasks.

Exhibit 6. Organizational Chart



AIR is uniquely qualified for this role, having successfully led four projects for IES in the last decade focused on teacher PD interventions. These projects involved coordinating across subcontracted organizations, including an intervention provider and several school districts, recruiting participating schools and teachers, and conducting an independent evaluation (see Exhibit 7). AIR’s experience monitoring intervention providers and providing feedback on fidelity for continuous improvement also helps ensure relevant, actionable feedback from the

evaluation team, which will draw on instruments and methods that AIR has refined across several studies. **AIR Lead Staff:** Andrew Wayne, Project Director (PD); Dorothy Seidel, Deputy Project Director (DPD); Marlene Darwin, Partnerships Lead; Mengli Song, Evaluation Lead (EL); Jane Coggshall, Evaluation Lead for Fidelity & Qualitative.

Exhibit 7. AIR-Led IES Contracts to Evaluate Teacher PD Interventions

Project Name	Subcontractor(s) That Provided the Intervention	Number of Districts and Schools
Focusing on Mathematical Knowledge: The Impact of Content-Intensive Teacher PD	Intel Math (summer institute), Mathematics Learning Community along with district-based coaches (school-year meetings), Harvard University along with district-based coaches (video-based coaching)	6 districts 73 schools
The Impact of Providing Performance Feedback to Teachers and Principals	Danielson Group, Teachscape, University of Virginia, Discovery Education (performance feedback)	8 districts 127 schools
Middle School Mathematics Professional Development Impact Study	America’s Choice & Pearson Achievement Solutions (summer institute, school-year meetings, coaching)	12 districts 77 schools
The Impact of Two Professional Development Interventions on Early Reading Instruction and Achievement	Language Essentials for Teachers of Reading and Spelling (summer institute/school-year meetings), district-based staff trained by the Consortium on Reading Excellence (coaching)	6 districts 90 schools

The Danielson Group and Learning Forward, co-designers of PLI, will provide all the necessary supervision, training, tools, and support coaches need to implement PLI. This includes all of the efforts described in Section B.2, Strategy to Implement PLI at Scale With High Quality. The Danielson Group’s lead coach (LC, Pamela Rosa) will emphasize delivery and refinement of workshops on the FfT Clusters and the instructional coaching model, while Learning Forward’s lead facilitator (LF, Michelle Bowman King) will emphasize preparation of the PLI coaches for their roles facilitating the Teacher Learning Teams. The LC and LF will co-lead each of the ongoing support sessions, including the group and individual support sessions.

Educopia’s project director (Eric Docter) will be responsible for the PLI online system. Educopia is well suited for this work, as it has in-depth experience creating and supporting secure, customizable online platforms for states and others to assess teaching quality incorporating multiple video-based and other measures of practice (see Educopia.com).

In addition to its role in delivery and refinement of PLI, Learning Forward will (1) lead outreach to identify any additional districts needed for the project and (2) convene meetings of the school district partners and others to support and sustain work of participating districts and other districts with interest in PLI. These activities leverage Learning Forward’s unparalleled state, regional, and national networks of educators, central office leaders, consultants, and others focused on improving teacher PD (see learningforward.org).

The school district partners will be supported by the other partner organizations to (1) recruit and select PLI coaches, (2) implement PLI, and (3) participate in all data collections (see Exhibit 8).

Exhibit 8. Group Responsible, Time Frame, and Milestones for Each Strategy

Milestones	Responsible	Project Year (October 1–September 30)				
		Year 1	Year 2	Year 3	Year 4	Year 5
Objective 1. Implement strategy to scale while continuously using feedback and fidelity data for project improvement.						
Strategy 1.1	Danielson Group LC, Districts	√	√	√		
Strategy 1.2	Districts	√	√	√		
Strategy 1.3	Danielson Group LC, Learning Forward LF	√	√	√	√	
Strategy 1.4	Danielson Group LC, Learning Forward LF		√	√	√	√
Strategy 1.5	Danielson Group LC, Learning Forward LF	√	√	√	√	√
Objective 2. Implement PLI while continuously using feedback and fidelity data for project improvement.						
Strategy 2.1	Danielson Group LC, Learning Forward LF	√	√	√	√	
Strategy 2.2	Districts		√	√	√	√
Strategy 2.3	Districts		√	√	√	√

Milestones	Responsible	Project Year (October 1–September 30)				
		Year 1	Year 2	Year 3	Year 4	Year 5
Strategy 2.4	Danielson Group LC, Learning Forward LF		√	√	√	√
Objective 3. Conduct an implementation evaluation (all cohorts) and RCT to test the impact of PLI on the quality of instruction and student outcomes (Cohorts 2 and 3 only).						
Strategy 3.1	AIR Evaluation lead, Districts	√	√	√		
Strategy 3.2	AIR Evaluation lead		√	√		
Strategy 3.3	AIR Evaluation lead, Fidelity and Qualitative lead		√	√	√	√
Strategy 3.4	AIR Evaluation lead		√	√	√	
Strategy 3.5	AIR Evaluation lead			√	√	√
Objective 4. Develop sample and infrastructure for continued scaling of PLI.						
Strategy 4.1	Learning Forward LF; Danielson Group LC; AIR Partnerships lead			√		
Strategy 4.2	Learning Forward Project Director; AIR Partnerships lead		√	√	√	√

C.3. Procedures Ensuring Feedback and Continuous Improvement

The project is designed to ensure feedback and continuous improvement through (1) its sequenced cohort structure and (2) routines for using feedback. Each cohort’s participation will create feedback that informs real-time improvements or improvements for the next cohort. The routines for gathering feedback and deciding on improvements are integrated into the meetings planned for the operation of the project (see Exhibit 9). Each meeting will include a regular agenda item **to discuss feedback and implications for improving the project’s materials, strategies, and procedures.** For example, during a *Monthly Implementation Review Meeting* just after the launch of Cohort 1, the LC and LF may report that all PLI coaches attended the *Monthly Group Support Check-Ins* (part of Strategy 1.4) but only half were engaged. The review team could decide that the LC and LF should use the *Monthly Individual Support Check-In* meetings to clarify expectations and ask each disengaged coach about making the group check-in meetings more engaging. The LC

and LF could then try modifying the meeting content or format as suggested. If that is successful, the review team would discuss implications for future instances of the *Biweekly Coaching Quality Meetings* and for materials for future implementations.

Exhibit 9. Routines for Using Feedback

Meeting Name and Frequency (When Active)	Participants	Feedback Data Sources	Strategies to be Improved
Monthly implementation reviews	AIR: PD, DPD, Partnerships Lead Danielson Group: LC Learning Forward: LF	All fidelity data collections (see Appendix G)	1.1, 1.2, 1.3, 1.4, 2.1, 2.2, 2.3
Monthly group support check-ins	Danielson Group: LC Learning Forward: LF PLI coaches	Coach interviews; coach logs; PLI FOI rubrics; number of coaching rounds and team cycles completed	1.4., 2.2, 2.3
Monthly individual support check-ins	Danielson Group: LC Learning Forward: LF PLI coaches	PLI FOI rubrics; number of rounds and team cycles completed	1.4, 2.2, 2.3
Monthly partnering network team meetings	Learning Forward: PL Danielson Group: LC AIR: Partnerships Lead	Partner engagement records	4.1, 4.2
Weekly evaluation team meetings	AIR: EL and Evaluation Team	Data collection and analysis update memo	3.1, 3.2, 3.3, 3.4, 3.5
Bimonthly evaluation reviews	AIR: EL and PD, Vice President	Data collection and analysis update memo	3.1, 3.2, 3.3, 3.4, 3.5
Monthly district check-ins (separately by district)	AIR Evaluation Lead, District Point-of-Contact	Data collection update	3.1, 3.2, 3.3, 3.4, 3.5
Biannual Danielson Group and Learning Forward design team meetings	Danielson Group: LC Learning Forward: LF	All fidelity data collections (see Appendix G)	1.1, 1.2, 1.3, 1.4, 2.1, 2.2, 2.3

C.4. Ongoing Work Beyond the End of the Grant

The project’s benefits will continue into the future through the partners, who each see the work as aligned to their mission. The Danielson Group, Learning Forward, and Educopia want the PLI program and all its services to be effective, affordable, and widely used. These three partners plan to use the project’s final materials and procedures in future implementations of PLI. For the

Danielson Group, the project also fits with its ongoing efforts to refine the FfT and apply it in powerful ways. Learning Forward’s mission includes helping states and districts identify scalable, effective PD that can be integrated into a coherent system of teacher support. Learning Forward plans to continue convening users and potential users of PLI to discuss experiences with the program and expanding its use. The school district partners will have trained, experienced, local coaches and resources earmarked for teacher PD that can support ongoing delivery of PLI. In addition, the cadre of coaches and teachers in each school district partner, trained through the project, have the potential to have an impact on new classes of students every year. The school district partners also will be well-positioned to consider expanding the use of PLI, either to additional schools or additional grade levels in the schools that are already participating.

To maximize impact during and after the project, AIR plans to seek foundation support for delivery of PLI via the scaling strategy to the control teachers in each cohort, delayed 2 years from the treatment group. AIR also plans to identify broader lessons from the evaluation to inform the development and enhancement of other PD programs. Finally, AIR will continue to pursue opportunities to build partnerships to scale, refine, and test teacher PD for high-need students, consistent with its mission to conduct and apply the best possible research toward improving people’s lives with an emphasis on the disadvantaged.

D. Quality of the Project Evaluation

AIR will conduct an independent evaluation to answer seven research questions (RQs) about the impact and implementation of the PLI program: (RQ1) What is the impact of PLI on teachers’ self-efficacy for improving classroom practice and the quality of teachers’ classroom practice? (RQ2) What is the impact of PLI on student engagement and academic achievement? (RQ3) To what extent is the impact of PLI on the quality of teachers’ classroom practice moderated by

teacher/classroom and school characteristics? (RQ4) To what extent is the impact of PLI on student achievement moderated by student, teacher/classroom, and school characteristics? (RQ5) To what extent is the impact of PLI on student achievement mediated by the quality of teachers' classroom practice? (RQ6) To what extent is PLI implemented with fidelity? (RQ7) What are the factors that hinder or facilitate the implementation of PLI?

These RQs will be addressed with data collected from three successive cohorts of schools as shown in Exhibit 1. The first cohort will include six schools that will implement PLI as a pilot cohort and provide data to address RQs 6 and 7 (i.e., the implementation questions). The second and third cohorts will include 72 schools in total (36 treatment and 36 control), which together will provide sufficient statistical power to detect the impact of PLI on key student and teacher outcomes (see Appendix G.2 for details about the power analysis) and contribute to answering all seven RQs.

All schools participating in the evaluation will be regular elementary schools serving high proportions of high-need students (i.e., 75% or more students eligible for free or reduced-price lunch), each with three or more regular Grade 4 teachers. The schools must already set aside time for Grade 4 teachers to meet weekly as a team, or be willing to do so, and teachers must not be departmentalized. Teachers in both treatment and control schools will be subject to their districts' normal PD requirements and opportunities,¹ but Grade 4 teachers in treatment schools also will participate in PLI for 2 years.

¹ The only exception is that in treatment schools where PLCs exist already, districts' normal PLC activities for Grade 4 teachers will be replaced by the facilitated TLT activities as part of the PLI program.

D.1. Evaluation Methods Designed to Meet WWC Evidence Standards

Without Reservations

Evaluation of the impact of PLI will be based on a blocked cluster RCT in which schools in the second and third cohorts will be randomly assigned to the treatment and control conditions within each of the six partner districts.² For this evaluation, schools are the appropriate unit of assignment because the TLTs—a key component of PLI—are school-based. Based on our prior experience with school-level RCTs, we expect minimal school-level attrition over the 2 years of the intervention.³ Given that the proposed evaluation is based on a school-level RCT that is free of confounding factors and is expected to have low attrition, it will produce strong evidence about the impact of PLI that will likely meet the WWC evidence standards without reservations.

D.2. Generation of Guidance About Effective Strategies Suitable for Replication

The proposed evaluation will generate useful guidance about effective strategies for implementing and scaling PLI in diverse settings by (1) including a large sample representing diverse settings; (2) deliberately assessing whether the impact of PLI differs for different types of students, teachers, classrooms, and schools; (3) collecting and analyzing rich data on program implementation from multiple sources; and (4) including a cost analysis to provide valuable information about the cost-effectiveness of the program.

Diverse Settings. The commitment of six partner districts that include a large number of high-need schools in diverse settings (See Section B.1) will allow the evaluation to generate valuable

² For large districts, we may further group schools into multiple random assignment blocks based on school characteristics that may be associated with student achievement, which will likely improve the statistical power of the impact estimates.

³ An IES-funded school-level RCT on the impact of providing teachers and principals with performance feedback recently completed by AIR (Garet et al., 2017), for example, included 127 schools in total, and only 1 school dropped out of the study during the 2-year intervention.

guidance for future replications of PLI in a variety of settings. (See Appendix D for letters of support and Appendix G.3 for the demographic characteristics of the partner districts.)

Differential Impact Analyses. The evaluation will include differential impact analyses (RQs 3 and 4) to assess the extent to which PLI’s impact is moderated by the characteristics of students, teachers/classrooms, and schools (see Exhibit 10). Results from these exploratory analyses will be crucial in guiding future efforts to scale PLI, as they may identify settings and populations for which the program is particularly effective or not well suited.

Exhibit 10. Potential Moderators at the Student, Teacher/Classroom, and School Levels

Student-Level Moderators	Teacher/Classroom-Level Moderators	School-Level Moderators
Race/ethnicity, eligibility for free or reduced-price lunch, English language learner status, special education status, and prior achievement scores	Teacher experience, probationary status, class size, and classroom average prior achievement	School size and demographic composition (e.g., percentage of minority students/ students from low-income families), and cohort (Cohort 2 versus 3)

Analyses of Implementation Data From Multiple Sources. To provide lessons learned for future replications or testing of PLI in other settings, the evaluation team will collect and analyze rich implementation data from multiple sources collected from the treatment schools from all three cohorts. In addition to implementation-related information tracked by the PLI online system (e.g., level of participation), we will examine implementation fidelity (RQ6) based on data from coaching logs, TLT logs, coach interviews, LC and LF interviews, implementation fidelity rubrics completed for a random sample of coaching cycles and TLT cycles, and teacher surveys.⁴ (See Appendix G.4 for details about implementation data collections.) Using these sources and the other measures identified in Section C.1, we will examine the fidelity of the implementation of both the scaling strategy (e.g., the selection of PLI coaches and coach

⁴ The teacher survey will be administered to both treatment and control teachers each spring, which will allow us to gather data on control teachers’ coaching experience as well to assess “service contrast.”

training, monitoring, and support) and the intervention itself (e.g., the dosage and the quality of instructional coaching and TLT activities), and identify factors that facilitate or hinder the implementation of the scaling strategy and the intervention (RQ7).

Cost Analysis. To provide information about whether PLI is a cost-effective investment and identify ways to make it more cost effective, we will conduct a cost analysis using the Resource Cost Model (RCM), which has been used extensively by AIR.⁵ Focusing on both personnel and nonpersonnel resources used in PLI, we will populate the RCM using the *CostOut* tool and generate cost-effectiveness estimates based on the cost estimates and results from the impact analyses.⁶

D.3. Valid and Reliable Performance Data on Relevant Outcomes

Teacher Outcomes: Classroom Practice and Self-Efficacy. According to the theory of change presented in Exhibit 2, the primary teacher outcome for PLI is the quality of teachers' classroom practice. For each teacher in the RCT sample, we plan to video-record one lesson in the early fall of the first intervention year (as baseline) and two lessons in the spring of the second intervention year. Certified FfT observers at AIR will code the videos, blind to condition. A subset (10%) of the lessons will be double-coded by independent coders to assess reliability. In addition to our primary measure, the FfT, we will use parallel procedures to code video-recorded lessons from the second intervention year using the Classroom Assessment and Scoring System-Upper Elementary (CLASS-UE)—to check the robustness of findings about PLI's impact on teacher outcomes not as directly aligned with the intervention. Similar to the FfT, CLASS-UE can be used across subjects and focuses on several dimensions of classroom practice. Both instruments

⁵ See <http://www.air.org/topic/p-12-education-and-social-development/school-finance>.

⁶ The *CostOut* tool is a tool created by the Center for Benefit-Cost Studies in Education at Teachers College, Columbia University.

have rich evidence of reliability and validity (Bill & Melinda Gates Foundation, 2012; Goe, Bell, & Little, 2008). Consistent with prior research, we will use the FfT overall score based on 10 observable dimensions.⁷ (See Appendix G.5 for dimensions measured by FfT and CLASS-UE.).

In addition to classroom practice, we also plan to assess the impact of PLI on a key intermediate teacher outcome as specified in PLI's theory of change—teachers' self-efficacy for improving classroom practice.⁸ For this outcome, we will use the *Teachers' Sense of Self-Efficacy Scale* (short form) developed by Tschannen-Moran and Woolfolk-Hoy (2001).⁹ It includes three subscales: *Instructional Practices*, *Classroom Management*, and *Student Engagement*, which are well aligned with the key constructs in the FfT that is the basis of PLI and have sufficient reliability (alpha = 0.81~0.86 for the three subscales and 0.90 for the overall scale). We will gather baseline data on teachers' self-efficacy with a short survey administered in the early fall of the first intervention year to all teachers in the RCT sample, and will include the self-efficacy measures in the spring teacher survey conducted in each intervention year.

Student Outcomes: Engagement and Achievement. As shown in the theory of change (Exhibit 2), key student outcomes for PLI are engagement and academic achievement. We will measure student engagement with a short survey administered in both the fall (as baseline) and spring of each intervention year. The survey will measure the quality and nature of students' participation in learning activities based on measures developed by Skinner, Furrer, Marchand, and Kindermann (2008) and Skinner, Kindermann, and Furrer (2009). It will include items associated with four engagement scales with alpha reliability ranging from 0.70 to 0.79 (see

⁷ We use the original FfT because it has well established measurement characteristics and aligns with the Clusters.

⁸ The theory of change identifies some additional intermediate teacher outcomes—teacher knowledge and skill. We do not measure these because no reliable measures of knowledge and skill related to the FfT are available.

⁹ See <https://www.statisticssolutions.com/teachers-sense-of-efficacy-scale-tses/> for details about the measure.

Appendix G.6). Measures of student achievement will be based on students' scores on state tests in mathematics and ELA. (See Appendix G.7 for technical details about all impact analyses.)

D.4. Clear Articulation of Components, Mediators, and Outcomes and Measurable Threshold

The design of the proposed evaluation is informed by clearly articulated key components, mediators, and outcomes of PLI as depicted in the theory of change presented in Exhibit 2. As Exhibit 2 shows, the central components of the PLI program include preparatory teacher workshops, instructional coaching, and TLT collaborative inquiry cycles informed by the FFT Clusters. The theory of change also specifies intermediate teacher outcomes (i.e., knowledge and skills, and self-efficacy for improving instructional practices and student engagement) and key outcomes for teachers (i.e., quality of classroom practices) and students (i.e., engagement and achievement). The intermediate teacher outcomes mediate PLI's impact on teachers' classroom practice, which in turn mediates the program's impact on student outcomes.

The evaluation specifies measurable thresholds for acceptable implementation of the PLI scaling strategy (described in Section B.2) and for acceptable implementation of the PLI program itself (described in Section A.3). For the former, a PLI coach must complete 95% of the six days of in-person coach training and participate in at least two thirds of the biweekly individual support meetings and at least two thirds of the biweekly group support meetings. For the PLI program itself, (1) a teacher must complete 85% of the 3-day Preparatory Workshop and complete at least 75% of the 16 planned coaching cycles; and (2) a coach must demonstrate successful implementation of at least 80% of the key elements of PLI based on the PLI Fidelity of Implementation Rubrics completed for a randomly selected sample of coaching cycles. These thresholds will be used to assess implementation fidelity and inform continuous improvement.

References

- Abt Associates. (2015). National evaluation of i3: NEi3 criteria and procedures for assessing i3 local evaluations excerpt: Independence. Cambridge, MA: Author.
- Agodini, R., Harris, B., Thomas, M., Murphy, R., & Gallagher, L. (2010). *Achievement effects of four early elementary school math curricula: Findings for first and second graders* (NCEE 2011-4001), Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Allen, J. P., Hafen, C. A., Gregory, A. C., Mikami, A. Y., & Pianta, R. (2015). Enhancing secondary school instruction and student achievement: Replication and extension of the My Teaching Partner-Secondary intervention. *Journal of Research on Educational Effectiveness*, 8(4), 475–489.
- Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science*, 333(6045), 1034–1037.
- Angrist, J. D., Imbens, G. W., & Rubin, D. (1996). Identification of causal effects using instrumental variables. *Journal of American Statistical Association*, 91, 444–455.
- Auguste, B. G., Hancock, B., & Laboissiere, M. (2009). *The economic cost of the U.S. education gap*. McKinsey & Company. Retrieved from <https://www.mckinsey.com/industries/social-sector/our-insights/the-economic-cost-of-the-us-education-gap>
- Baker, D. W., Wolf, M. S., & Feinglass, J. (2007). *Health literacy and mortality among elderly persons*. *Archives of Internal Medicine*, 167(14), 1503–1509. Retrieved from <https://jamanetwork.com/journals/jamainternalmedicine/article-abstract/412862>

- Bill & Melinda Gates Foundation. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains*. Seattle, WA: Author.
- Blazar, D., Kraft, M. A., & Hogan, D. (2017). *The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence*. Paper presented at the spring 2017 conference of the Society for Research on Educational Effectiveness, Washington, DC.
- Borman, G. D., Dowling, N. M., & Schneck, C. (2008). A multisite cluster randomized field trial of Open Court Reading. *Educational Evaluation and Policy Analysis*, 30(4), 389–407.
- Chamberlain, G. E. (2013). Predictive effects of teachers and schools on test scores, college attendance, and earnings. *Proceedings of the National Academy of Sciences*, 110(43), 17176–17182. Retrieved from <http://www.pnas.org/content/110/43/17176.full>
- Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of Teachers II: Teacher value-added and student outcome in adulthood. *American Economic Review*, 104, 2633–2679.
- Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of Teachers II: Teacher value-added and student outcome in adulthood. *American Economic Review*, 104, 2633–2679.
- Chingos, M. M., & Whitehurst, G. J. (2012). *Choosing blindly: Instructional materials, teacher effectiveness, and the Common Core*. Washington, DC: The Brown Center on Education Policy at Brookings. Retrieved from https://www.brookings.edu/wp-content/uploads/2016/06/0410_curriculum_chingos_whitehurst.pdf
- Cohen, D., & Ball, D. L. (1999). *Instruction, capacity, and improvement*. Philadelphia, PA: Consortium for Policy Research in Education.
http://www.cpre.org/sites/default/files/researchreport/783_rr43.pdf

- Dong, N., & Maynard, R. A. (2013). PowerUp!: A tool for calculating minimum detectable effect sizes and sample size requirements for experimental and quasi-experimental designs. *Journal of Research on Educational Effectiveness*, 6(1), 24–67.
- Cowan, J., Goldhaber, D., & Theobald, R. (2017). *Teacher equity gaps in Massachusetts* [ESE Policy Brief]. Malden, MA: Massachusetts Department of Elementary and Secondary Education. Retrieved from <http://www.doe.mass.edu/research/reports/2017/10teacher-equity.pdf>
- Danielson, C. (2008). *The handbook for enhancing professional practice: Using the framework for teaching in your school*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, N., & Orphanos, S. (2009). *Professional learning in the learning profession: A status report on teacher development in the United States and abroad*. Dallas, TX: National Staff Development Council.
- DeLuca, C., Shulha, J., Luhanga, U., Shulha, L. M., Christou, T. M., & Klinger, D. A. (2015). Collaborative inquiry as a professional learning structure for educators: a scoping review. *Professional Development in Education*, 41(4), 640–670.
- Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis* 24(2): 81–112.
- Desimone, L. M., & Garet, M. S. (2015). Best practices in teacher's professional development in the United States. *Psychology, Society, & Education*, 7(3), 252–263.

- Desimone, L., Smith, T. M., & Phillips, K. J. R. (2013). Linking student achievement growth to professional development participation and changes in instruction: A longitudinal study of elementary students and teachers in Title I schools. *Teachers College Record*, 115(5), 1–46.
- Dynarz, K. (2018). Checking in: Are math assignments measuring up? Washington, DC: The Education Trust. Retrieved from https://1k9gl1yevnfp2lpq1dhrqe17-wpengine.netdna-ssl.com/wp-content/uploads/2014/09/CheckingIn_MATH-ANALYSIS_FINAL_5.pdf
- Easton, L. B. (Ed.). (2008). *Powerful designs for professional learning*. (2nd ed.). Oxford, OH: National Staff Development Council.
- Edmondson, A. C., & Roloff, K. S. (2009). Overcoming barriers to collaboration: Psychological safety and learning in diverse teams. *Team Effectiveness in Complex Organizations: Cross-Disciplinary Perspectives and Approaches*, 34, 183–208.
- Ermeling, B. A. (2009). Tracing the effects of teacher inquiry on classroom practice. *Teaching and Teacher Education*, 26(3), 377–388. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0742051X09000559>
- Gabriel, R., Peiria Day, J., & Allington, R. (2011). Exemplary teacher voices on their own development. *Phi Delta Kappan*, 92(8), 37–41.
- Garet, M. S., Heppen, J. B., Walters, K., Parkinson, J., Smith, T. M., Song, M., ... Borman, G. D. (2016). *Focusing on mathematical knowledge: The impact of content-intensive teacher professional development* (NCEE 2016-4010). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

- Garet, M. S., Heppen, J. B., Walters, K., Smith, T. M., & Yang, R. (2016). *Does content-focused teacher professional development work? Findings from three Institute of Education Sciences studies* (NCEE 2017-4010). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Retrieved from <https://ies.ed.gov/ncee/pubs/20174010/pdf/20174010.pdf>
- Garet, M., Porter, A., Desimone, L., Birman, B., & Yoon, K. (2001). What makes professional development effective? Analysis of a national sample of teachers. *American Education Research Journal*, 38(3), 915–945.
- Garet, M. S., Wayne, A. J., Brown, S., Rickles, J., Song, M., & Manzeske, D. (2017). *The impact of providing performance feedback to teachers and principals* (NCEE 2018-4001). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from <https://ies.ed.gov/ncee/pubs/20184001/pdf/20184001.pdf>
- Garrett, R., & Steinberg, M. (2015). Examining teacher effectiveness using classroom observation scores: Evidence from the randomization of teachers to students. *Educational Evaluation and Policy Analysis*, 37, 224–242.
- Gennetian, L. A., Morris, P. A., Bos, J. M., & Bloom, H. S. (2005). Constructing instrumental variables from experimental data to explore how treatments produce effects. In H. S. Bloom (Ed.), *Learning more from social experiments: Evolving analytic approaches* (pp. 75–114). New York, NY: Russell Sage Foundation.
- Gershenson, S. (2016). Linking teacher quality, student attendance, and student achievement. *Education Finance and Policy*, 11(2), 125–149.

- Gersten, R., Taylor, M. J., Keys, T. D., Rolhus, E., & Newman-Gonchar, R. (2014). *Summary of research on the effectiveness of math professional development approaches* (REL 2014–010). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Education Research Laboratory Southeast. Retrieved from http://ies.ed.gov/ncee/edlabs/regions/southeast/pdf/REL_2014010.pdf
- Goe, L., Bell, C., & Little, O. (2008). *Approaches to Evaluating Teacher Effectiveness: A Research Synthesis*. National Comprehensive Center for Teacher Quality.
- Goldhaber, D., Lavery, L., & Theobald, R. (2015). Uneven playing field? Assessing the teacher quality gap between advantaged and disadvantaged students. *Educational Researcher*, 44(5), 293–307.
- Goldhaber, D., Quince, V., & Theobald, R. (2016). *Reconciling different estimates of teacher quality gaps based on value-added*. Washington, DC: National Center for the Analysis of Longitudinal Data in Education Research (CALDER), American Institutes for Research.
- Hasiotis, D., Jacobs, A., & McGovern, K. (2015). *The mirage: Confronting the hard truth about our quest for teacher development*. Washington, DC: The New Teacher Project. Retrieved from <https://tntp.org/publications/view/evaluation-and-development/the-mirage-confronting-the-truth-about-our-quest-for-teacher-development>
- Hedberg, E. C., & Hedges, L. V. (2014). Reference values of within-district intraclass correlations of academic achievement by district characteristics: Results from a meta-analysis of district-specific values. *Evaluation Review*, 38(6), 546–582.
- Hirsh, S., & Crow, T. (2017). *Becoming a learning team: A guide to a teacher-led cycle of continuous improvement*, Oxford, OH: Learning Forward.

- Hirsh, S. (2018). *Focus professional learning communities on curriculum* [blog post]. *Education Week Teacher*. Retrieved from http://blogs.edweek.org/edweek/learning_forwards_pd_watch/2018/01/focus_professional_learning_communities_on_curriculum.html
- Isenberg, E., Max, J., Gleason, P., Johnson, M., Deutsch, J., & Hansen, M. (2016). Do low-income students have equal access to effective teachers? Evidence from 26 districts (NCEE 2017-4007). Washington, DC: Institute of Education Sciences.
- Jacob, R. Zhu, P., & Bloom, S. H. (2009). *New empirical evidence for the design of group randomized trials in education*. New York, NY: MDRC.
- Jensen, B., Sonnemann, J., Roberts-Hull, K., & Hunter, A. (2016). *Beyond PD: Teacher professional learning in high-performing systems*. Washington, DC: National Center on Education and the Economy. Retrieved from <http://www.ncee.org/wp-content/uploads/2015/08/BeyondPDWeb.pdf>
- Kane, T. J., & Staiger, D. O. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains*. Seattle, WA: Bill & Melinda Gates Foundation. Retrieved from <https://files.eric.ed.gov/fulltext/ED540960.pdf>
- Kane, T., Owens, A., Marinell, W., Thal, D., & Staiger, D. (2016, February). [*Teaching higher: Educators' perspectives on Common Core implementation*](#). Cambridge, MA: Center for Education Policy Research, Harvard University.
- Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*.
- Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, 86(4), 945–980.

- Lash, A., Tran, L., & Huang, M. (2016). *Examining the validity of ratings from a classroom observation instrument for use in a district's teacher evaluation system* (REL 2016–135). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West. Retrieved from <https://files.eric.ed.gov/fulltext/ED565904.pdf>
- Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). New York, NY: Wiley-Interscience.
- Lomos, C., Hofman, R. H., & Bosker, R. J. (2011). Professional communities and student achievement—A meta-analysis. *School Effectiveness and School Improvement*, 22(2), 121–148.
- NAEP. (2017a). *No changes in fourth-grade mathematics scores for most student groups compared to 2015. NAEP mathematics report card: National student group scores and score gaps. The nation's report card*. Retrieved from https://www.nationsreportcard.gov/math_2017/#/nation/gaps?grade=4
- NAEP. (2017b). *No significant change in the percentage of fourth-grade students at or above Proficient in reading compared to 2015. NAEP reading report card: National achievements-level results. The nation's report card*. Retrieved from https://www.nationsreportcard.gov/reading_2017/#/nation/achievement?grade=4
- NAEP. (2018). *2017 NAEP mathematics and reading assessments: Highlighted results at grades 4 and 8 for the nation, states, and districts. The nation's report card*. Retrieved from https://www.nationsreportcard.gov/reading_math_2017_highlights/

- National Center for Education Statistics. (n.d.). Table 5. Number and percentage of public schools that had staff with specialist or academic coaching assignments, and among those schools, the percentage that had staff with a particular specialist or academic coaching assignment, by selected school characteristics. *National Teacher and Principal Survey*. Retrieved from https://nces.ed.gov/surveys/ntps/tables/Table_5_042617_fl_school.asp
- National Center for Education Statistics. (2017). *The condition of education 2017*. Washington, DC: Institute of Education Sciences. Retrieved from <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2017144>
- Neufeld, B., & Donaldson, M. L. (2012). Coaching for instructional improvement: Conditions and strategies that matter. In B. Kelly & D. F. Perkins (Eds.), *The Cambridge handbook of implementation science for educational psychology* (pp. 373-92). Cambridge, UK: Cambridge University Press.
- New Teacher Center. (2016). *High quality mentoring and induction practices*. Chicago, IL: Author. Retrieved from https://newteachercenter.org/wp-content/uploads/high-quality-mentoring_induction-resource.pdf
- Olson, L. (2018). *High-quality curricula and team-based professional learning: A perfect partnership for equity*. Dallas, TX: Learning Forward. Retrieved from <https://learningforward.org/wp-content/uploads/2018/05/curriculaPLEquity.pdf>
- Opfer, V.D., Kaufman, J.H., & Thompson, L.E. (2016). *Implementation of K–12 state standards for mathematics and English language arts and literacy*. Santa Monica, CA: RAND Corp.
- Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods* 15(3), 209–233.

- Raghunathan, T. E., Lepkowski, J. M., Van Hoewyk, J., & Solenberger, P. (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey Methodology*, 27(1), 85–95.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2), 417–458.
- Rotermund, S., DeRoche, J., & Ottem, R. (2017). *Teacher professional development by selected teacher and school characteristics: 2011–12*. Washington, D.C.: National Center for Education Statistics, Institute for Education Sciences, U.S. Department of Education. Retrieved from <https://nces.ed.gov/pubs2017/2017200.pdf>
- Sartain, L., Stoelinga, S. R., & Brown, E. R. (2011). *Rethinking teacher evaluation in Chicago: Lessons learned from classroom observations, principal-teacher conferences, and district implementation. Research Report*. Chicago, IL: Consortium on Chicago School Research.
- Sass, T. R., Hannaway, J., Xu, Z., Figlio, D. N., & Feng, L. (2012). Value added of teachers in high-poverty schools and lower poverty schools. *Journal of Urban Economics*, 72(2), 104–122.
- Saunders, W. M., Goldenberg, C. N., & Gallimore, R. (2009). Increasing achievement by focusing grade-level teams on improving classroom learning: A prospective, quasi-experimental study of Title I schools. *American Educational Research Journal*, 46(4), 1006–1033.
- Schochet, P. Z., & Chiang, H. (2009). *Technical methods report: Estimation and identification of the complier average causal effect parameter in education RCTs* (NCEE 2009-4040). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

- Selig, J. P., & Preacher, K. J. (2008, June). *Monte Carlo method for assessing mediation: An interactive tool for creating confidence intervals for indirect effects* [Computer software]. Available at <http://quantpsy.org/calc.htm>
- Skinner, E., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of Educational Psychology, 100*(4), 765–781.
- <http://doi.org/10.1037/a0012840> Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection conceptualization and assessment of children’s behavioral and emotional participation in academic activities in the classroom. *Educational and Psychological Measurement, 69*(3), 493–525.
- Steinberg, M., & Sartain, L. (2015). Does teacher evaluation improve school performance? Experimental evidence from Chicago’s Excellence in Teaching Project. *Education Finance and Policy, 10*(4), 535–572.
- Steinberg, M. P., & Sartain, L. (2015). Does teacher evaluation improve school performance? Experimental evidence from Chicago’s Excellence in Teaching Project. *Education Finance and Policy, 10*(4), 535–572.
- Steiner, D. (2017). *Curriculum research: What we know and where we need to go*. Retrieved from <https://standardswork.org/wp-content/uploads/2017/03/sw-curriculum-research-report-fnl.pdf>
- Taylor, E., & Tyler, J. (2012). The effect of evaluation on teacher performance. *American Economic Review, 102*(7), 3628–3651.
- Tschannen-Moran, M., & Woolfolk-Hoy, A. (2001). Teacher efficacy: Capturing and elusive construct. *Teaching and Teacher Education, 17*, 783–805.

- The Education Trust. (2017). *Achievement and Opportunity in America: Where are We? What Can We Do?* Retrieved from <https://edtrust.org/resource/achievement-opportunity-america-can-2/>
- U.S. Department of Education. (2016). *The nation's report card: 2015 mathematics & reading at Grade 12*. Washington, DC: Author.
- U.S. Department of Education. (2014). "Fiscal year 2014 budget summary and background information." Washington, DC: Author. Retrieved from <http://www2.ed.gov/about/overview/budget/budget14/summary/14summary.pdf>
- U.S. Department of Education, Office of Innovation and Improvement. (n.d.). *Evaluation activities and responsibilities* (January 2015). Washington, DC: Author.
- Vangrieken, K., Dochy, F., Raes, E., & Kyndt, E. (2015). Teacher collaboration: A systematic review. *Educational Research Review*, 15, 17–40.
- Wayne, A. J., Yoon, K. W., Zhu, P., Cronen, S., & Garet, M. S. (2008). Experimenting with teacher professional development: Motives and methods. *Educational Researcher*, 37, 469–279.
- What Works Clearinghouse. (2017). *What Works Clearinghouse Standards Handbook* (Version 4.0). Washington, DC: Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf
- Wiener, R., & Pimentel, S. (2017). *Practice what you teach: Connecting curriculum and professional learning in schools*. Washington, DC: Aspen Institute. Retrieved from <https://assets.aspeninstitute.org/content/uploads/2017/04/Practice-What-You-Teach.pdf>
- Woulfin, S. L., & Rigby, J. G. (2017). Coaching for coherence: How instructional coaches lead change in the evaluation era. *Educational Researcher*, 46(6), 323–328.