Project Narrative (Mid-Phase)

Introduction

The effectiveness of early career teachers (ECTs) is a crucial lever for elevating student achievement, especially in high-poverty schools. This proposal exemplifies Absolute Priorities 1 and 4. It builds on evidence that the instructional coaching cycle designed by Teachstone supports students' social and emotional learning (Collaborative for Academic, Social, and Emotional Learning, 2023) and has outsized impacts on ECTs, advancing their students' achievement by 2 months or more (see A4). The focal program is called "Positive Acceleration for Early Career Teachers (+AFFECT)." It combines the optimal dosage of Teachstone coaching cycles with group activities and leadership support, creating a scalable and sustainable program likely to make ECTs in high-poverty schools much more effective.

The American Institutes for Research® (AIR®) will lead the overall project and ensure coordination across partners. Teachstone will implement +AFFECT in collaboration with the six district partners. AIR and Georgia State University (GSU) will conduct the independent evaluation of +AFFECT, producing findings that inform not only Teachstone and the district partners but also a national audience of education practitioners, policy makers, and researchers. Learning Forward (LF) will lead the partnership's collaborative dissemination efforts.

In total, the project will refine, test, scale, and sustain the +AFFECT program in 76 high-poverty middle schools, divided into three cohorts (see Exhibit 1). The first cohort is a pilot cohort. Cohorts 2 and 3 will be the basis for a multisite school-level randomized controlled trial (RCT). AIR will recruit 28 and 44 schools, respectively, for these two impact cohorts, randomly assigning half of each cohort's schools to treatment and half to delayed treatment (control).

This multicohort design creates opportunities to refine and sustain +AFFECT. Across the arc of the project, Teachstone will provide +AFFECT directly for 1 year in each pilot cohort school

and each treatment school in Cohorts 2 and 3, iteratively improving it using data and feedback. Teachstone will then train and support local district staff, who will extend +AFFECT to the delayed-treatment schools and encourage all schools to sustain the use of the program.

Exhibit 1. Number of Schools That Will Receive the Program, by Cohort and Year

School cohort	2024–25	2025–26	2026–27	2027–28
Cohort 1 (pilot cohort)	4 pilot schools			
Cohort 2 (impact cohort)		14 treatment schools	14 control schools (delayed treatment)	
Cohort 3 (impact cohort)			22 treatment schools	22 control schools (delayed treatment)

CPP1: Partners to Support Equity. GSU's team includes two co-principal investigators (co-PIs) and a graduate assistant, all focused on the evaluation (see B2). AIR and GSU have been working together for 4 years to promote diversity in the education sciences and to bring diverse perspectives and lived experiences into research activities. We continue to deepen our institutional relationship, collaborating on projects and proposals. For this project, GSU's team brings complementary research skills and relationships with high-poverty schools across the Atlanta-area, including in one of the partner districts (Fulton). GSU's role exemplifies CPP1 because GSU is a minority-serving institution and is not typically an EIR grantee.

CPP2. The project specifically addresses the impact of COVID-19 on students and teachers in high-need schools (see A1). For **CPP2a**, it uses student engagement as a means to improve student achievement (see A2), and its individual coaching incorporates regular, individual needs assessment by trained instructional coaches (see A3). For **CPP2b**, +AFFECT supports engaging, effective instruction using an evidence-based instructional approach (see A2) and an evidence-based approach to instructional coaching (see A3, A4), all without tracking or remedial courses.

A. Significance

A1. ECTs in High-Poverty Schools: A Promising Lever for Equity in Education

Studies that examine variation in student outcomes all point to **teachers as a strong lever** for improving outcomes, including achievement, social and emotional competencies, and later life outcomes (e.g., Chamberlain, 2013; Chetty et al., 2014; Gershenson, 2016; Kraft, 2019; Rivkin et al., 2005; Sass et al., 2010). In a study that randomly assigned students to teachers to eliminate the possibility that "teacher effects" reflect schools' sorting of students into classes, researchers found that having a teacher at the 75th rather than the 25th percentile of teacher effectiveness translated to a big difference in achievement (+0.35 SD in reading; +0.48 SD in math). The authors summarized, "**Teacher effects are much larger than school effects**." In addition, variation in teacher effects was largest among the sample's high-poverty schools, meaning **teacher effects are even stronger in high-poverty schools than others** (Nye et al., 2004).

These findings are key as we respond to the drop in student achievement caused by COVID-19. The drop hit high-poverty schools the hardest and enlarged historic educational disparities. According to National Assessment of Educational Progress (NAEP) data (2023a, 2023b), the share of middle school students proficient or better in math was only 13% among low-income students, compared to 38% for others. In reading, the corresponding figures were 19% and 41%.

But the damage is deeper because **COVID-19 created teacher workforce problems in high-poverty schools that will have long-term consequences.** We know novice teachers are less effective than veteran teachers, yet novice teachers are assigned students with lower initial achievement (Center for Education Policy Research, 2012). In addition, even before the pandemic, the percentage of novice teachers in the workforce was about twice as high in high-poverty schools than in affluent schools (Carver-Thomas & Darling-Hammond, 2017). The

pandemic deepened this challenge, as all schools, but especially high-poverty schools, had to hire applicants they would not have hired pracademic (Delarosa & Elias, 2022). Unless labor market conditions change radically, student achievement in high-poverty schools will be limited by less-effective novices, who may then become less-effective veterans, or quit.

To raise achievement in high-poverty schools, we need programs that boost effectiveness for ECTs. This project is significant because it focuses on **students in high-poverty schools** assigned to **ECT classrooms**. We define high-poverty schools as Title I schools. We define ECTs as those with 5 years of experience or less. **ECTs in high-poverty schools are a high-leverage and critical focus for equity in education.**

The project's innovative, evidence-based strategy to address this need is +AFFECT, a 1-year teacher professional learning (PL) program for ECTs designed to be effective and scalable in high-poverty schools. As shown next, +AFFECT is grounded in a rubric designed for students from low-income families that is used by instructional coaches and teachers to **foster classroom interactions that support student engagement** (see A2). The main "active ingredient" in +AFFECT is Teachstone's virtual, individualized coaching cycle, which focuses on teacher-student and student-to-student interactions (see A3). With support and feedback from a trained coach, teachers complete five cycles, which minimizes burden and optimizes effectiveness, according to a recent RCT (see A4). To magnify impact and enhance scalability, +AFFECT incorporates streamlined group activities for the ECTs and leaders in each school (see B1).

A2. Student Engagement: A Promising Focus for ECTs in High-Poverty Schools

Teachers say that student engagement is a major barrier to learning recovery *and* a top priority for PL. Teachers surveyed by *Education Week* (2022) cited "problems with student engagement" more often than any other challenge to learning recovery. Eighty percent of

respondents said the pandemic made students less motivated in school (Prothero, 2023). And teachers' top priority for PL was "building student engagement and motivation" (Klein, 2021).

The +AFFECT program focuses ECTs on student engagement as a goal with multiple dimensions—emotional, behavioral, and cognitive engagement. ECTs often struggle with student engagement and fall back on behavioral compliance. But students learn best when their engagement is supported emotionally and cognitively as well. For example, students need teachers to foster a positive climate and respond productively to social-emotional needs. The focus of +AFFECT on authentic engagement builds on a consensus among researchers that authentic engagement in school drives academic learning (Reeve, 2013; Roorda et al., 2017; Skinner et al., 2009) and reduces future problems, such as dropout, depression, and aggressive behaviors (Li & Lerner, 2011; Samelo-Aro, 2017; Voisin & Elsaesser, 2016).

To support engagement, the PL activities in +AFFECT are guided by a system for observing features of classroom interactions that predict student engagement and learning during academic instruction in any subject ([redacted] et al., 2013; Pianta, 2016). **These teacher-student interactions and student-to-student interactions are observable and "coachable."** Specifically, coaches who use +AFFECT learn an observation system called "Classroom Assessment and Scoring System-Secondary" (CLASS-S), originally developed to work in Head Start centers that serve children from low-income families and now adapted for adolescent students. CLASS-S defines the quality of classroom interactions using a validated 11-dimension rubric spanning three domains: Emotional Support, Classroom Organization, and Instructional Support. For example, one dimension within the first domain is Positive Climate (see Appendix J.1.1 for additional dimension descriptions and research citations). For each dimension, raters choose among seven distinct levels of performance. A recent statewide study linked a 1-point

increase on the 7-point CLASS-S overall scale to 4.4 and 5.1 percentile points of student growth in English language arts (ELA) and math achievement, respectively (Therriault et al., 2020). The use of a validated framework—and focus on classroom interactions—helps **make +AFFECT a unique and promising alternative to mainstream coaching programs** (e.g., Knight, 2017).

By design, the dimensions of observable interactions used to guide +AFFECT are especially relevant in high-poverty schools, where students bring diverse needs, assets, and interests culturally, linguistically, and academically. Teachers using +AFFECT learn to interact in ways that are responsive to these and other important factors, such as students' need for autonomy and decision making with sufficient support to achieve a sense of control (Allen et al., 1994; Anderman & Midgley, 1998). CLASS-S dimensions—notably Positive Climate and Regard for Adolescent Perspectives—overlap with parameters for two of the four principles of culturally responsive teaching defined by Wlodkowski and Ginsburg (1995a, 1995b) (see also Reese et al., 2014). And as we describe later (see C3), students in high-poverty schools receive the types of student engagement support described by CLASS-S much less frequently than other students. +AFFECT will help teachers in high-poverty schools boost engagement and achievement.

A3. Instructional Coaching: Teachstone's Unique Approach

Teachstone's coaching cycle is a promising alternative to traditional coaching, based on how the program (a) focuses on classroom interactions (see A2); (b) is mediated through an online system; (c) specifies clear, practical steps for coaches and teachers; (d) centers on video clips, enabling virtual rather than in-person learning; and (e) uses positive reinforcement only, rather than an "areas for growth" approach. At the start of each cycle, the teacher uploads a video recording of a typical lesson, and the coach finds three clips that each illustrate how a classroom interaction supported student engagement (see Exhibit 2 and detail in Appendices J.1.2 & J.1.3).

The cycle requires the teacher to reflect on those clips in writing and then orally with the coach. These steps help the teacher understand vividly what drove engagement and why, and how they can use supportive interactions to achieve their teaching goals. Coaches and teachers praise the focus on "strengths-based" positive reinforcement (Foster, 2021; Wells & Foster, 2022).

Exhibit 2. The Five Steps of Teachstone's Coaching Cycle

Step 1. The teacher uploads a video capturing 30+ minutes of classroom instruction.

Step 2. The coach views the video and selects three 1-minute clips that illustrate teacher-student interactions and engagement. For each clip, the coach writes a question to spur reflection.

Step 3. The teacher views the video and questions and posts responses, which the coach reads before the videoconference.

Step 4. During a 20- to 30-minute videoconference, the teacher and coach discuss the video and the teacher's responses to the prompts. The coach guides the teacher in reflection, helping the teacher understand how the quality of classroom interactions affects student engagement. The two also discuss goals for the teacher, focusing on particular dimensions of classroom interactions, and develop an action plan.

Step 5. The coach sends the teacher a detailed written conference summary and action plan.

Teachstone's individual coaching cycle is also promising because of its **feasibility for implementation at scale**. Each step is asynchronous except for the videoconference, making it a convenient alternative to traditional coaching (Foster, 2021). And a cycle takes a teacher just 1–2 hours, spread across 2–3 weeks (Clark et al., 2022). To ensure fidelity, Teachstone staff called "specialists" have protocols, materials, and tools that comprise the Fidelity Support System (see Appendix J.1.4), which they use to train coaches, monitor their work, and intervene with coaches individually when needed. The cycle is used widely in Head Start centers to serve pre-K teachers with disadvantaged students but is not yet well known in K–12 classrooms.

In general, coaching is a promising alternative to traditional PL approaches such as workshops or summer institutes because it individualizes PL, grounds it in each teacher's classroom practice, gives teachers agency, and establishes ongoing accountability between the teacher and the coach (Hill & Papay, 2022). A meta-analysis found positive results for students on average, though not all programs produced positive results (Kraft et al., 2018).

A4. Building on Findings From Successive Impact Evaluations

The project design draws on lessons from three prior impact evaluations of Teachstone's coaching model, which reveal three key insights that motivate the project and its design.

First, the individual coaching was effective in demonstrating positive impacts on student achievement **in varied settings and for diverse students**. The program produced consistent positive impacts in two trials in distinct settings: in middle schools in predominantly white, rural/small town, middle-income districts (Allen et al., 2011), and in high schools in a predominantly black, large urban, high-poverty district (Allen et al., 2015; see Evidence Form for details on the two trials). Impacts did not vary by student characteristics (race and ethnicity, poverty, gender, prior achievement). Moreover, follow-up analysis of the first trial showed reduced use of exclusionary discipline on Black students (Gregory et al., 2013a, 2013b, 2015).

Second, a **lower dosage of coaching produced better results**. The third and biggest trial involved 14 large urban and suburban districts. Among the students, two thirds were from low-income families. The sample enabled the researchers to test two different dosage levels of the individual coaching: five cycles and eight cycles. The five-cycle program resulted in a positive impact equivalent to 2 months of learning on ELA achievement, while the eight-cycle version did not produce a statistically significant impact on achievement (Clark et al., 2022). The authors hypothesize that in the eight-cycle version, teachers did not have sufficient time between cycles to incorporate what they learned about classroom interactions and student engagement.

The third insight motivating this proposal's design is that **the most promising population for the individual coaching is ECTs**. Among teachers with 0–5 years of experience, the five-cycle program had a positive impact across **both ELA and math** (2.5 months of learning in math *and* 2.75 months in ELA; Clark et al., 2022). ECTs may be more responsive to PL

experiences or to particular features of Teachstone's coaching cycle.

Summary. ECTs are a critical lever to improve achievement in high-poverty schools. To accelerate their development and make a lasting strategic impact on achievement and the teacher workforce in high-poverty schools, the project advances +AFFECT—a feasible, convenient program that is well suited to high-poverty schools, promises months of additional learning, works across subject areas, and fills a need for PL focused on student engagement. The project will refine, test, scale, and sustain +AFFECT. As discussed next in B1 and depicted in the logic model (see Appendix G), the +AFFECT program includes individual coaching (Teachstone's coaching cycle) and complementary school leader and ECT team activities that support scaling.

B. Strategy to Scale

B1. Strategies to Overcome Past Barriers to Scaling and Sustaining

To ensure that all schools implement the +AFFECT program for 1 year and that many sustain it, the project plans four key strategies to overcome past barriers to scaling and sustaining.

Strategy 1. Alignment With the Felt Needs of District and School Leaders. In a recent implementation, the largest barrier to scale was the initial buy-in of district and school leaders. We asked them to sign up for Teachstone coaching cycles at a dosage of 10 cycles per year for 2 years, and many declined quickly. The decision to limit +AFFECT to five cycles—and to focus on ECTs—aligns not only with impact evidence but also with the felt needs of district and school leaders revealed in dialogue when writing this grant application: (a) especially since COVID-19, leaders are careful about adding activities for teachers; (b) ECTs are a "pain point"—principals struggle to find ways to help ECTs become more effective without overwhelming them; and (c) as federal recovery funds end, leaders seek strategies that are sustainable and cost-effective. Cost analyses show the five-cycle dosage is highly cost-effective (Clark et al., 2022).

Districts representing 170 Title I middle schools signed letters of support (6) to try out +AFFECT for their ECTs—far more than needed to fill Cohorts 1, 2, and 3 (see Appendix J.3). We will engage additional districts as needed by leveraging the networks of AIR, GSU, LF, and Teachstone. An additional partner, Teach for America, will introduce +AFFECT to its network of district clients with high-poverty schools to support more scaling.

Strategy 2. Program Activities Supporting Scaling—School Leadership Activities. Teachstone's prior implementations of the individual coaching lacked mechanisms to engage leadership—important supports for scaling and for teacher PL program effectiveness (Grissom et al., 2021; Learning Forward, 2022). Principals said they lacked opportunities to receive updates (e.g., the number of completed cycles), set priorities (e.g., particular domains of classroom interaction), or provide other input. Involving principals and other instructional leaders would have enabled them to intentionally sustain the program and support further scaling.

To address this barrier, +AFFECT includes school leadership activities. Specifically, Teachstone's assigned coach will hold four meetings across the year with each school's leadership, to include the principal and a principal-appointed instructional leader (e.g., a coach or mentor) designated the "ECT champion," who also plays a role in ECT team activities (see Strategy 3). The ECT champion role will be supported by Teachstone's Fidelity Support System (see Appendix J.1.4) via an initial training and one-on-one check-ins between the four leadership meetings (see Appendix J.1.5 for the ECT champion manual's table of contents).

At the first school leadership meeting, the principal and ECT champion will self-assess and share other contextual information (e.g., school culture, priorities) to inform planning and steer program implementation. At the remaining three meetings, Teachstone, the principal, and the ECT champion will discuss implementation, troubleshoot issues (e.g., coordination with existing

initiatives affecting ECTs), and eventually plan for the continued integration of +AFFECT practices in the next year (see Strategy 4). These activities will engage school leadership, improve PL coherence and integration (Darling-Hammond et al., 2016; Desimone, 2009), and lay groundwork for continued implementation after the first year.

Strategy 3: Program Activities Supporting Scaling—ECT Team Activities. Another barrier to scaling and sustaining is the negative stigma teachers feel about being selected for coaching. Instructional coaches often focus on teachers who are struggling, so teachers assume receiving coaching is a sign of deficiency (Mangin & Dunsmore, 2014; Picucci & Laughlin, 2019). In prior implementations, Teachstone carried out the individual coaching without any activities to ensure that participation had a positive social meaning. Teachers were reluctant to discuss their coaching experiences with others, even though the experiences were positive (e.g., Foster, 2021) and sharing those experiences would have supported scaling.

To address this barrier, +AFFECT includes ECT team activities. The activities are designed to ensure that ECTs and others see the program as a supportive, collective initiative that includes all ECTs. To that end, the ECT champion and Teachstone coach will cofacilitate a fall and spring team meeting in each participating school. For the meeting, each ECT will select a favorite video clip from their coaching cycles to share and discuss with the group. The principal will join the spring team meeting to stay informed and to celebrate the ECTs' efforts. Teachstone piloted this strategy and found that teachers used the strategy in team meeting times with fidelity (Stuhlman et al., 2022). Such collaborative PL activities should broaden and deepen buy-in for the program, help it self-propagate, and potentially improve its effectiveness (Garet et al., 2001; Kraft, 2019).

Strategy 4: Support for Continued Integration After the First Year. Beyond school leadership and teacher buy-in, scaling and sustaining a reform requires creating "conditions to

shift authority and knowledge of the reform from external actors to teachers, schools, and district" (Coburn, 2003). Our **specific strategy** to overcome this barrier—the need to shift authority and knowledge—is to give schools and central office leaders greater (a) control over program parameters and then (b) knowledge and responsibility for the coaching.

Toward the end of the first year of implementation, AIR will facilitate dialogue among Teachstone, district leadership, and school leadership and a sample of ECTs at schools finishing their first year of implementation. To shift authority, central office and school leadership will decide program parameters for the second year, such as which teachers receive coaching, how many coaching cycles, and which dimensions of classroom interactions to emphasize.

During the second year, Teachstone will continue to engage school leadership and support team activities remotely but will shift responsibility for individual coaching to the district. Using guidance from Teachstone to select suitable staff (see B3 and Appendix J.2), the central office will allocate the time of one or more district coaches. Teachstone will train, monitor, and support the district coaches using its Fidelity Support System (see Appendix J.1.4). Teachstone uses this system to support its own coaches and has shown it successfully produces the individual coaching with fidelity using local district coaches ([redacted] & Coggshall, 2022). District coaches will serve treatment schools after the first year of program implementation. They will also serve teachers in the 36 delayed treatment schools and could serve additional schools.

B2. Management Plan to Achieve the Objectives on Time and Within Budget

Appendix J.4 depicts the broad responsibilities of each partner. Along the top row, AIR's *project management* team manages the project. The *evaluation* team includes AIR and GSU staff, who will conduct the independent evaluation of the +AFFECT program.[1] Along the bottom

[1] Consistent with Office of Elementary and Secondary Education guidance for independent evaluation (Abt Associates, 2020),

row, Teachstone and the school districts implement +AFFECT, and LF leads dissemination.

As the management plan chart in Appendix J.6 shows, AIR assigns responsibility for each project objective and specific strategy (all identified in C2) to a specific partner *and* lead staff member. The chart also specifies each strategy's timeline and milestones. To achieve these project objectives and strategies on time and within budget, AIR and the partners each bring a proven history of implementing and scaling projects as well as well-qualified personnel (see B3).

To implement the management plan, AIR's project director ([redacted]) and deputy director ([redacted]) will hold a project kickoff meeting for each project objective with the responsible staff to review and revise management plan, then hold recurring task-based meetings. AIR will monitor progress and costs; ensure coordination across partner organizations and school districts; and adjust the plan as needed to ensure timeliness and performance within the project budget.

B3. Organizational Capacity to Bring the Project to Scale

The partnership depicted in Appendix J.4 provides the management capacity and qualified personnel needed to bring the project to scale. (For staff resumes, see Appendix B.) Using its management infrastructure, AIR has a distinguished track record of projects to scale up and evaluate PL programs in high-need schools. In each of the three example projects shown in Appendix J.7, AIR coordinated across subcontracted organizations, including a PL provider and several school districts; recruited participating districts, schools, and teachers; and conducted an independent evaluation that informed program improvements and assessed impacts.

AIR's management plan designates experienced staff for lead roles, including its project management team and evaluation team. Leading AIR's project management team, [redacted]

the evaluation team is separate from the project management team and will have no role in the development or implementation of +AFFECT 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 enforced through corporate conflict of interest policies, team-specific charge codes, and bimonthly review by senior leaders in AIR.

[redacted], ***project director,*** brings 20 years of experience studying teachers and teaching, including leading several projects that refine, test, and scale teacher PL programs designed to improve outcomes in Title I or other high-need schools. [redacted], ***deputy director,*** has 10 years of project management experience and has successfully managed large-scale projects that focus on teacher PL in high-need schools. Before coming to AIR, she taught as a bilingual teacher in a Title I school. [redacted], ***partnerships lead,*** is an experienced educator, principal, and superintendent who works with Title I schools regularly on AIR's district and school improvement team and has built and managed partnerships on two teacher PL projects.

The evaluation team, which spans AIR and GSU, also brings experienced staff. [redacted], ***evaluation lead,*** has designed and executed studies of interventions including teacher PL programs for 13 years, including in schools with students from marginalized racial and ethnic groups and from low-income families. [redacted], ***fidelity and qualitative lead*** just completed a parallel role on an EIR grant focused on PL with Danielson Group. At GSU, [redacted], ***GSU co-PI,*** is on the faculty in the Dept. of Learning Sciences. An HBCU graduate, he conducts research on secondary education, including deeper learning and special education. He will contribute to instrument development, data collection, and dissemination together with a GSU graduate assistant. [redacted], ***GSU co-PI*****,** is a senior research scientist in the College of Education and Human Development and has expertise in large evaluation studies. She will leverage her relationships with local education agencies to support recruitment and data collection.

Teachstone will provide all supervision, training, tools, and support needed to implement +AFFECT. Founded in 2008, Teachstone is the national vendor for +AFFECT and several other PL and assessment services to support effective classroom interactions. The organization recently delivered its individual coaching with fidelity in an RCT involving 107 schools across

14 districts (Clark et al., 2022). [redacted]***, program design lead,*** will lead updates to the design of +AFFECT, drawing on her expertise in teacher PL, classroom processes, and serving students in disadvantaged schools. [redacted]***, program implementation lead,*** will lead Teachstone's delivery of services to schools. [redacted] brings more than 20 years of experience and has successfully directed and managed large projects in education, working with high-need schools and examining culturally responsive instructional practices. Teachstone's trained and CLASS-S certified specialists, such as [redacted], will oversee a team of **experienced, professional coaches**, with the following qualifications: CLASS-S certification; experience teaching and providing professional development to education professionals; and proven skill in virtual training and facilitation (for coach qualifications, see Appendix J.2).

LF will lead the dissemination work (see B4), drawing on its leadership and capacity to inform, reach, and engage audiences interested in educator PL (see *learningforward.org*). [redacted]***, dissemination lead,*** will direct this work and engage partners and participants, including the community of practice (CoP) (see B4). As LF's vice president for publications, [redacted] oversees the planning, creation, production, and dissemination of all LF products, including member newsletters, *The Learning Professional* journal, and books.

The **six school district partners** bring leaders and educators committed to strong implementation of PL that contributes to student success. Each superintendent will appoint a district leader (usually an asst. superintendent) to collaborate with the project team to recruit and support Title I schools to participate. To scale and sustain +AFFECT after the first year, the leader will recruit district coaches using Teachstone's guidelines for coach qualities (Appendix J.2). Each principal will ask a current instructional leader (e.g., teacher leader, instructional coach) with experience motivating adults and teams to serve as the school's ECT champion.

B4. Dissemination Mechanisms to Support Further Development and Replication

The dissemination will support (a) broad awareness of project information, program practices, and lessons; and (b) district and school replication of the program and key program practices in diverse contexts. LF will lead collaborative efforts to author open license products such as a communication toolkit with infographics, PowerPoint presentations, video testimonials, blogs, and articles (see B5; for a list of products, see Appendix J.9). LF will leverage (a) each partner's dissemination mechanisms, (b) third-party mechanisms, and (c) a CoP, as detailed here. The partners will use social media to publicize the products, which will be available at no cost.

LF Dissemination Mechanisms. LF is a widely trusted broker of meaningful information and dialogue about PL among K–12 professionals. Its members include more than 32,000 central office and building administrators, teachers, PL providers, and researchers. The project team and participants will author dissemination products for its journal (*The Learning Professional)*, its blogs on learningforward.org, its "white paper" series as well as webinars and annual conference sessions. Products will feature participants' experiences, innovative program practices, lessons from implementation, and findings about impact. Social media outreach and announcements in LF's weekly, monthly, and quarterly newsletters (*PD in the News*, *Connect*, and *Tools for Learning Schools*) will draw attention to products, all available for free on the project webpage.

Teachstone Dissemination Mechanisms. Teachstone will leverage its strong market presence and website as well as its annual InterAct conference, typically attended by hundreds of participants who seek to learn how to better support effective classroom interactions. These audiences will hear stories about +AFFECT experiences, program features, and lessons to bring to their home districts. To support replication of the program and its key practices, Teachstone will publicly post products designed to support replication (i.e., manuals and other resources

identified in Appendix J.9 and B5) and offer cost-based training services.

AIR Dissemination Mechanisms. AIR brings reputation and reach across a swath of practitioners, policy makers, and researchers who seek to use the best evidence. AIR will support replication and development by featuring no-cost products on its website, which logs hundreds of thousands of visits monthly. AIR will also use its email networks that target those who use evidence to support high-quality teaching—for example, the dissemination networks of its Center on Great Teachers and Leaders and of Regional Educational Laboratories, each of which engage practitioners, policy makers, and researchers. Accessible products featuring the program design, evidence base, and project lessons will engage these audiences broadly.

GSU Dissemination Mechanisms. In addition to its reach among researchers and policy makers nationally, GSU partners with school districts across Georgia and supplies a significant share of the state's teachers. To encourage replication, GSU will engage these partners in dialogue about the program and participant experiences as well as project stories and lessons.

Third-Party Dissemination Mechanisms. AIR and GSU will reach researchers through peer-reviewed academic journals and at research conferences. The partners will also develop broadly understandable products like highlights, FAQs, and briefs for education trade publications (e.g., *Education Week*, *Educational Leadership*) and conferences (e.g., Association for Supervision and Curriculum Development; Education Commission of the States).

CoP Dissemination Mechanisms. Within partner districts, LF will recruit participating ECTs, ECT champions, principals, and district administrators to join the CoP and convene them three times per year, growing the CoP across cohorts. CoP meetings will elicit in-depth discussion of participants' needs and experiences with +AFFECT, which will inform program adjustments (see bottom of C3). CoP members will also discuss expanded use of +AFFECT in

participating districts and share strategies for addressing barriers to scale and sustain the program. The CoP will thus support replication in additional schools and districts.

B5. Useful Products Tailored to be Effective for End Users

The free information, materials, and processes we develop and disseminate through the mechanisms described in B4 will support three target audiences: (1) policy makers, PL provider organizations, and school system staff whose support is needed; (2) school system staff who decide to implement the program or its key techniques and practices; and (3) researchers.

For the first group, the most useful products will be short blogs, articles, conference presentations, and webinars, using outlets that these audiences trust (see examples in Appendix J.9). In our experience, these products are most effective when they incorporate the voices of program participants from diverse settings, so we will invite varied participants to author products or share their perspectives with authors at AIR, Teachstone, and LF. Participants will also receive a Communications Toolkit to support sharing their perspectives and project information with other districts as well as regional and national stakeholders.

Second, for districts and schools that decide +AFFECT can help them, the project will disseminate products that support replication of the program or key program practices and techniques. To that end, refined implementation guides (for central office staff and school leaders), detailed manuals (including coach, teacher, and ECT champion manuals), and FAQs will be posted online. These open-license materials will incorporate refinements based on Teachstone's continuous improvement efforts, which will draw on its operational data as well as evaluation feedback from AIR and the CoP. Materials will include tips and guidance for successful use in various settings, based on the project's diverse sample. With these materials, local coaches anywhere can try the coaching techniques. Teachstone will also continue to offer

+AFFECT services directly as well as cost-based services to train, monitor, and support local coaches who can implement the program using the fidelity support system.

Finally, for researchers, we will present at research conferences and publish a report and journal article on study findings, which will inform future research projects and innovations.

C. Quality of the Project Design

C1. Clearly Articulated Conceptual Framework Underlying the Project

Underlying the project activities is a coherent conceptual framework, shown in Appendix J.5. The base depicts *Teachstone resources supporting program implementation*, including Teachstone's *Fidelity Support System*, *specialists*, and *coaches*. The +AFFECT program consists of *individual coaching* and *program activities supporting scaling.* These key program components are expected to help ECTs understand and identify effective interactions, and then incorporate them into their classroom practices, creating a positive impact on the primary teacher outcome: *quality of classroom interactions*. The use of more effective interactions should, in turn, lead to improved student outcomes: *student engagement* and *academic achievement*. Given its use of positive reinforcement, the program may also have positive impacts on two secondary teacher outcomes: *sense of self-efficacy* and *job satisfaction*.

This framework is grounded in evidence cited throughout this proposal. Teachstone's Fidelity Support System successfully yields coaching with fidelity, including in high-poverty schools ([redacted] & Coggshall, 2021), and its individual coaching has demonstrated positive impacts on student achievement, as well as on classroom interactions and engagement in high-poverty and diverse settings (based on citations in A4). +AFFECT's school leadership and ECT team activities are grounded in research on features of PL that make it effective (see citations in B1). A large body of evidence links the quality of classroom interactions, student engagement,

and student achievement for students in high-poverty schools (see A2 and Appendix J.1.1). And the Teachstone coaching cycle improves workforce outcomes similar to job satisfaction (see C3).

C2. Clearly Specified and Measurable Goals, Objectives, and Outcomes

AIR will use its management plan (see B2 and Appendix J.6) to ensure the partnership achieves the project's overall goal: to refine, test, scale, and sustain +AFFECT in 76 schools (4 + 28 + 44) between 2025 and 2028, and to disseminate lessons and practices that advance the field. The management plan breaks this goal into three objectives and 13 strategies, each with specific, measurable outcomes, as shown in Exhibit 3. An objective's strategies repeat across years, as detailed in the table. For example, Objective 1 is to implement and improve +AFFECT. We will complete all the strategies under Objective 1 for each pilot school in Cohort 1, then each treatment school in Cohort 2, and then each treatment school in Cohort 3.

Exhibit 3. Objectives, Strategies, Outcomes, and Measures

Strategies	Measures	Outcomes
Objective 1: Implement +AFFECT for 1 year while regularly using feedback and fidelity data for project improvement (each school in Cohort 1 and each treatment school in Cohorts 2 and 3).		
Strategy 1.1 AIR recruits schools and teachers for pilot and impact study.	Signed memoranda of understanding (MOUs) for participating schools	4 schools sign MOUs for Cohort 1; 28 sign for Cohort 2; and 44 sign for Cohort 3.
Strategy 1.2 Teachstone recruits and trains coaches to provide PL to teachers.	Completion of program training and CLASS-S certification	2, 6, 15, and 9 coaches certify in CLASS-S and complete Teachstone's training for 2025–26, 2026–27, 2027–28, and 2028–29 school years respectively (includes delayed treatment schools).
Strategy 1.3 Teachstone implements +AFFECT program, including coaching and program activities supporting scaling.	Coaching logs and coaching session attendance records	Coaching logs and meeting records show 80% of teachers participate in four or more coaching cycles.
Strategy 1.4 AIR collects formative evaluation data from teachers, coaches, and coach specialists on teacher PL experiences and feasibility.	Teacher, coach, and coach specialist surveys; pilot study fidelity of implementation briefs	Survey responses show that 80% of teachers and coaches and 100% of coach specialists have completed feedback surveys.
Strategy 1.5.Teachstone refines program materials, including coach training and supports and activities related to coaching.	All measures in Appendix J.10 (e.g., Teachstone online system, coach logs, FOI checklist, coach-teacher relations survey, etc.).	Data and feedback show improved coaching supports and fidelity of implementation during a given cohort and from one cohort to the next.

Strategies	Measures	Outcomes
Objective 2: Measure +AFFECT's impact on teacher and student outcomes (Cohorts 2 and 3).		
Strategy 2.1 AIR randomly assigns schools to treatment and control groups.	Treatment–control balance in baseline measures of the outcomes and other school characteristics	Random assignment results in treatment–control balance in baseline measures of the outcomes and other school characteristics, with no baseline group differences exceeding 0.25 SD.
Strategy 2.2 AIR collects and analyzes fidelity data and shares with Teachstone.	Coach logs, implementation briefs	90% of teachers in treatment schools have completed all five coaching cycles.
Strategy 2.3. AIR collects teacher and student data.	Completion rates of teacher surveys, video recordings, and extant achievement data	All collections achieve response rates of 90% or greater.
Strategy 2.4. AIR determines service contrast and +AFFECT impact on teachers & students.	Classroom observations, teacher surveys, and state assessment data	Findings about the impact of +AFFECT on teacher and student outcomes (RQ1&2).
Objective 3: Disseminate project findings and replicate and sustain program use.		
Strategy 3.1. Teachstone launches and convenes community of practice (CoP).	Briefs from CoP convenings and CoP participant feedback	Increased engagement and capacity to sustain and expand +AFFECT in participating districts
Strategy 3.2. Learning Forward (LF) and partners create written and video products and disseminate them.	Number of products, article downloads, website clicks, video views, social media mentions, and mentions on other digital platforms	Increased awareness and use of +AFFECT and key program practices
Strategy 3.3. LF and partners present findings at research and practitioner conferences.	Number of conference presentations and participant feedback	Two conferences per year, in Years 3, 4, and 5, and increased knowledge among potential +AFFECT users and the field
Strategy 3.4. Teachstone supports delayed implementation schools in implementing +AFFECT through training, monitoring, and support for district coaches.	Training attendance records, meeting notes, coaching logs	Sustained use of +AFFECT in treatment schools and expanded use of +AFFECT beyond treatment schools in participating districts

C3. Project Design That Addresses the Needs of the Target Population

The target population is students assigned to ECTs in high-poverty middle schools. The project design addresses the needs of these students, their school leaders, and their ECTs.

Students. COVID-19 deepened challenges with student engagement (see A1), which is already a struggle when adolescents enter middle school (Eccles & Roeser, 2009). These students need engaging classroom experiences that promote learning, and +AFFECT helps teachers provide that. Its validated instructional framework is thoroughly responsive to the

diverse needs and interests of students in high-poverty schools (see A2). The partner districts' Title I middle schools reflects this diversity. Among students in these schools, on average, 76% qualify for lunch, 47% are Black, 30% are Hispanic, and 13% are English learners (Appendix J.3). The districts signed up because they see +AFFECT as a way to meet these students' needs.

The need of these students to receive engaging instruction is far from being met, according to the best evidence available. Data from an AIR study (Garet et al., 2017) that video-recorded classrooms in 127 schools in 8 districts show that **students in classrooms in the highest poverty schools experienced the lowest levels of support for engagement as measured by all three domains of CLASS-S**—especially Emotional Support and Classroom Organization. For example, among higher poverty schools, 39% of classrooms were in the lowest quartile for Emotional Support, compared to 14% in lower poverty schools, as shown in Appendix J.11.

Early Career Teachers. ECTs know they are learning and face an urgent need to engage students authentically and reliably. +AFFECT targets that need, and RCTs of the +AFFECT coaching cycle prove that it helps ECTs learn and grow in their effectiveness. In fact, the impact of the coaching has been greater for ECTs than veteran teachers (see A4). Of teachers who experienced the Teachstone coaching cycle, 97% reported it to be relevant ("it is relevant to my instruction"), 88% to be specific ("it provided specific ideas about how I could improve"), and 94% to be effective in improving their knowledge ("it changed how I think about my instruction") ([redacted] et al., 2023). Teachers from high-need schools in geographically diverse settings describe the coaching features that make it useful to them in a YouTube video (https://tinyurl.com/MTP-Video) and in a digital story (https://tinyurl.com/MTP-story), both emphasizing its simplicity, strengths-based approach, and use of video clips.

School Leaders. These leaders face a persistent challenge: ECTs are generally less effective

than veterans, but the share of teachers who are ECTs is comparatively higher in high-poverty schools (see A1). The project design addresses these school leaders' felt need to support ECTs with PL that is low-burden, effective, and sustainable (see A4, B1). The project prioritizes middle schools because the need for PL about student engagement is strongest there, as PL programs for secondary teachers are usually content focused (Kraft & Blazar, 2018).

School leaders also seek ways to encourage ECTs and to deter quitting or switching schools. A national study found annual teacher turnover was 50% higher in Title I schools compared to elsewhere (Carver-Thomas & Darling-Hammond, 2017). +AFFECT helps address this challenge by building self-efficacy (see A3). In fact, a recent trial of Teachstone's coaching model in schools where 64% of students were lunch-eligible found a large, positive impact (0.77 SD) on a construct called "teachers' enthusiasm about teaching" ([redacted] et al., 2023).

Ongoing Need-Sensing and Adjustment. Finally, the project design facilitates regular need-based adjustments based on multiple mechanisms. Teachstone refines +AFFECT regularly based on implementation experiences, AIR evaluation data (see D2), and input from the CoP (see B4), and adjusts to local needs. The coaching cycle itself uses video to adjust to individual teachers.

D. Quality of the Project Evaluation

D1. Generation of Evidence That Meets WWC Standards Without Reservations

The independent evaluation will be conducted by AIR in partnership with GSU (see B2 for details about ensuring the independence of the evaluation). The evaluation will answer seven research questions (RQs) about the impact and implementation of the +AFFECT program: (RQ1) What is the impact of +AFFECT on the quality of interactions in ECTs' classrooms and on ECTs' sense of self-efficacy and job satisfaction? (RQ2) What is the impact of +AFFECT on student engagement and academic achievement? (RQ3) To what extent is the impact of

+AFFECT on the quality of classroom interactions moderated by teacher and school characteristics? (RQ4) To what extent is the impact of +AFFECT on student engagement and achievement moderated by student, teacher, and school characteristics? (RQ5) To what extent is the impact of +AFFECT on student engagement and achievement mediated by the quality of classroom interactions? (RQ6) To what extent are the +AFFECT program components implemented with fidelity? (RQ7) What are the factors that hinder or facilitate the implementation of +AFFECT? We will address these questions using a rigorous evaluation design as described below, which will generate valuable information that will help inform the improvement and scaling of the +AFFECT program.

Evaluation Design. As Exhibit 1 shows, the evaluation is based on three successive cohorts of schools. Each cohort takes place in a different set of districts. The first cohort includes four schools that implement +AFFECT as a pilot cohort while the other two cohorts are impact cohorts. The four schools in the pilot cohort will provide initial implementation data to address RQs 6 and 7 and inform the refinement of the +AFFECT program for at-scale implementation in later cohorts. The 72 schools in the two impact cohorts will provide sufficient statistical power to detect an impact of +AFFECT on key teacher and student outcomes (see Appendix J.12 for power analysis details) and contribute the data needed to address all seven RQs.

For this RCT, school is the appropriate unit of assignment because +AFFECT is a school-level intervention that serves all ECTs in middle school grades and that incorporates school-level activities (i.e., school leadership activities and ECT team activities). AIR will randomly assign the 72 schools with equal probability to the treatment and control conditions within each study district. In large districts, we may form multiple random assignment blocks based on school characteristics associated with student achievement, to reduce chance imbalance and add power.

Each school will meet three eligibility criteria. First, each must be a public school serving at least two of the three middle school grades (Grades 6–8). Second, each must be eligible for Title I. Third, given the evaluation's focus on teacher instruction and student achievement in both math and ELA, each school also needs to have at least one math ECT and at least one ELA ECT in Grades 6–8 prior to random assignment at the beginning of the program year.

Teachers in all study schools will be subject to their district's normal professional development requirements and opportunities during the project period, but Grade 6–8 ECTs in treatment schools in each impact cohort also will participate in the +AFFECT program for 1 year, and their peers in control schools will be given the opportunity to participate in the program in the following year (delayed treatment).

Key Teacher and Student Outcomes. ELA and math achievement data are available for all middle school grades. To capitalize on these data, the primary student outcomes will come from students in the classroom of ECTs who teach ELA and math. We will also focus on these classrooms for observations to measure the primary teacher outcome—the quality of classroom interactions. This teacher outcome takes significant resources to measure correctly, and doing so in these classrooms will enable well-powered analyses of the extent to which classroom interactions mediate impacts on achievement (RQ5). We will gather other teacher outcomes from all ECTs using surveys, as these are easily gathered through teacher surveys.

To measure the quality of classroom interactions, we will conduct classroom observations of all Grade 6–8 math and ELA ECTs in our study sample and analyze the data separately by subject. Specifically, for each of these ECTs, we will video record one lesson in the early fall prior to the start of the treatment (coaching cycles) as baseline and three lessons at the end of the program year, with each lesson based on a randomly selected class section taught by the ECT.

All video-recorded lessons will be coded by certified coders at AIR who are blind to treatment status using both CLASS-S and Charlotte Danielson's Framework for Teaching (FFT) (Danielson, 1996), and a subset will be coded by all coders to check interrater reliability.[2] Both instruments measure the quality of classroom interactions and have rich evidence of reliability and validity (Bill & Melinda Gates Foundation, 2012; Pianta et al., 2012).[3] To assess +AFFECT's impact on the quality of classroom interactions (RQ1), we will analyze the CLASS-S overall score and its three domain scores as well as the FFT overall score and its two domain scores.[4] Given the potential overalignment issue with the CLASS-S measures, we will consider the FFT overall score the main measure of the quality of classroom interactions and consider the FFT domain scores and CLASS-S measures supplemental measures.

In addition to the primary teacher outcome, we plan to examine two secondary teacher outcomes—teachers' sense of self-efficacy and job satisfaction—measured with a survey administered to all ECTs (not just math and ELA ECTs) in study schools at both the beginning (baseline) and the end of the program year. The survey will use reliable survey scales from existing instruments (National Center for Education Statistics, 2012; Tschannen-Moran & Woolfolk Hoy, 2001) (see Appendix J.14 for the survey measures and their reliabilities).

The primary student outcomes for the evaluation are achievement in math and ELA, which will be analyzed separately based on standardized state test scores obtained from districts'

[2] While CLASS-S focuses exclusively on classroom interactions, FFT includes two domains focusing on classroom interactions and another two domains focusing on other professional skills. The FFT measures used for the proposed evaluation will be based on codings of the two FFT domains focusing on classroom interactions. The number of videos to be coded by all coders will depend on the number of videos collected and the number of coders available. We will estimate interrater reliability (correlation) using Stata's "icc" command based on variance composition.

[3] According to the CLASS-S manual (Pianta et al., 2012), the internal consistency reliabilities for the three CLASS-S domains range from 0.87 to 0.91. An EIR project recently completed by AIR researchers ([redacted] et al., 2023) achieved an interrater reliability of 0.86 for the CLASS-S overall score, 0.82 for the FFT overall score, and 0.52 to 0.85 for CLASS-S and FFT domain scores based on classroom observations. Further, the WWC has also accepted CLASS-S and FFT measures as teacher outcomes eligible for WWC review according to WWC's Data for Study Reviews database.

[4] Each score is an average across the relevant CLASS-S/FFT dimensions. See Appendix J.13 for domains and dimensions measured by CLASS-S and FFT and the creation of CLASS-S and FFT overall scores.

administrative records. As a supplemental student outcome, we will measure student engagement at the classroom level using the CLASS-S *Student Engagement* dimension score based on the classroom observations described earlier, which provide a more objective and cost-efficient way to measure student engagement than alternative measures such as self-reports. (See Appendix J.15 for technical details about multilevel analyses of teacher and student outcomes.)

Potential for Meeting WWC Standards Without Reservations. According to version 5.0 of What Works Clearinghouse (WWC) standards, a school-level RCT meets WWC standards without reservations if it has low attrition and limits the risk of bias due to joiners (What Works Clearinghouse, 2022, p. 43). To minimize attrition, we plan to conduct random assignment in the fall of the program year rather than the prior spring, thus eliminating potential attrition over the summer. That reduces teacher attrition to about one quarter of what NCES and others measure, which is year-over-year teacher mobility (Redding and Henry, 2017). For the same reason, it reduces student mobility and even reduces school attrition, as there is little time for changes in leadership or school priorities to occur. To further prevent attrition, we will (a) minimize time commitment and burden on study participants (see A3), (b) incorporate program features (e.g., school leadership activities and ECT team activities) that foster school buy-in and collective teacher participation, (c) offer data collection incentives to study teachers and delayed treatment to control schools to encourage continued participation, and (d) develop relationships with study schools and teachers to ensure continued cooperation with study activities.

To avoid the risk due to joiners, we plan to identify the teacher sample and student sample prior to random assignment; thus our impact analyses will not include any teacher joiners (e.g., replacement teachers for study teachers who leave their school before the end of the program year) or student joiners. This solution produces valid "intent-to-treat" impact estimates for

student and teacher outcomes and is consistent with WWC guidance. In addition, it eliminates the threat of cross-over, in which a teacher in one condition transfers to another.

Given that the evaluation is a school-level RCT that is likely to have low attrition and is free of confounding factors or risk of bias due to joiners and given that participant outcomes will be based on valid, reliable measures and consistent data collection procedures, impact findings for those outcomes clearly have the potential to meet WWC standards without reservations.

D2. Guidance About Effective Strategies Suitable for Replication or Testing

We will generate useful guidance about effective strategies to implement and scale +AFFECT in diverse settings, based on four features of the evaluation design.

Feature 1: Diverse Settings. The six partner districts contain 170 high-poverty schools potentially eligible for the evaluation. Located in diverse settings with respect to urbanicity and student demographics (see Appendix J.3), the diverse sample of schools will allow the evaluation to generate valuable guidance for replicating +AFFECT in a variety of settings.

Feature 2: Differential Impact Analyses. The evaluation will include additional analyses (RQs 3 and 4) to assess the extent to which +AFFECT's impact on key teacher and student outcomes is moderated by student, teacher, and school characteristics (see Exhibit 4). Findings about differential impact will help identify settings and populations for which +AFFECT is particularly effective or not well suited, and will help guide future efforts to scale the program.

Exhibit 4. Potential Moderators at the Student, Teacher, and School Levels

Student-level moderators	Teacher-level moderators	School-level moderators
Race/ethnicity, eligibility for free or reduced-price lunch, English learner status, special education status, and prior achievement	Whether a teacher is in their first year of teaching and whether a teacher has a graduate degree	School size, urbanicity, and demographic composition (e.g., percentage of minority students/ students from low-income families)

Feature 3: Analyses of Implementation Data from Multiple Sources. To provide lessons learned for future replications or testing of +AFFECT or similar programs, the evaluation team

will collect rich implementation data from multiple sources. These data will allow us to assess implementation fidelity (RQ6), identify factors that hinder or facilitate the implementation of +AFFECT (RQ7), and generate actionable feedback to inform Teachstone's iterative program improvement and scaling efforts as well as national discussions about teacher PL.

To assess implementation fidelity (RQ6), we will examine how each key program component was implemented with regard to dosage and quality. As detailed in Appendix J.10, the fidelity data we plan to collect include participation and fidelity records from trainings, teacher surveys,[5] and logs completed by ECT champions and Teachstone specialists and coaches. We also will collect records from the online system used by Teachstone to facilitate individual coaching cycles (see Appendix J.1.6). In addition to information about frequency and dosage, the online system maintains artifacts from coaching cycles (e.g., video clips, reflective questions, and session summaries). We will analyze these artifacts using the coaching fidelity checklist (see Appendix J.1.7) for two randomly selected coaching cycles per treatment school.

To further explore implementation-related issues, twice during the implementation year we will convene the ECT champions in each district for group interviews. We also will interview Teachstone specialists and coaches twice per year regarding their experience providing training and support. Qualitative analysis of these data and analysis of treatment teachers' responses to survey questions about their experience with +AFFECT will pay particular attention to factors that facilitate or hinder the implementation of the +AFFECT program (RQ7).

Feature 4: Cost and Cost-Effectiveness. To inform future replications of +AFFECT, we will analyze its cost and the cost-effectiveness, using the methods described in Appendix J.8.

[5] The teacher survey will be administered to teachers in both treatment and control schools, which will allow us to gather data on control teachers' coaching experience and to assess "service contrast." The survey for teachers in treatment schools will include a "treatment-only" section that asks about teachers' experience with +AFFECT with both Likert-scale and open-ended questions.

D3. Clear Articulation of Components, Mediators, Outcomes, and Thresholds

The conceptual framework clearly articulates key program components, mediators, and outcomes. As shown in Appendix J.5, the key components of the +AFFECT program include individual coaching as well as school leadership activities and ECT team activities designed to support the scaling of the program. The conceptual framework also specifies the primary teacher outcome (i.e., quality of classroom interactions) and secondary teacher outcomes (i.e., sense of self-efficacy and job satisfaction). The primary teacher outcome is hypothesized to mediate the program's impact on student outcomes (i.e., student engagement and academic achievement), which motivates RQ5. Appendix J.15 describes additional, exploratory mediation analyses.

We have specified measurable thresholds for acceptable implementation of key components of the +AFFECT program, drawing largely on Teachstone's experience delivering PL programs during the past 15 years. These thresholds are (a) Teachstone provides the initial orientation in all treatment schools, and at least 90% of the ECTs, on average, in treatment schools attend the virtual orientation or view a recording of it; (b) each ECT in treatment schools completes at least four of the five individual coaching cycles, on average, during the program year, and a randomly selected sample of coaching cycles demonstrates successful implementation of at least 80% of the key elements of the coaching cycles, based on the fidelity of implementation checklist (see Appendix J.1.7); (c) on average, at least three of the four planned meetings between the leadership (principal and ECT champion) of each treatment school and the Teachstone specialist occur during the program year; and (d) on average, at least 1.5 of the two planned meetings of ECTs in each treatment school occur during the program year with an average attendance rate of at least 80%. These thresholds will be used to address RQ6 regarding the fidelity of +AFFECT implementation in treatment schools and to inform continuous program improvement.

References

Abt Associates. (2020, March 26). EIR evaluation-related criteria and submission of draft EIR study design plans.

[redacted]

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 MyTeachingPartner-Secondary intervention. *Journal of Research on Educational Effectiveness*, *8*(4), 475–489. https://doi.org/10.1080/19345747.2015.1017680

Allen, J. P., Kuperminc, G., Philliber, S., & Herre, K. (1994). Programmatic prevention of adolescent problem behaviors: The role of autonomy, relatedness, and volunteer service in the Teen Outreach Program. *American Journal of Community Psychology, 22*(5), 617–638. https://doi.org/10.1007/BF02506896

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. https//doi.org/10.1126/science.1207998

Anderman, L. H., & Midgley, C. (1998). Motivation and Middle School Students. ERIC Digest.

Bill & Melinda Gates Foundation. (2012). Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains.

Bryk, A. S., Lee, V. E. & Holland, P. B. (1993). *Catholic schools and the common good.* Harvard University Press.

Carver-Thomas, D., & Darling-Hammond, L. (2017). *Teacher turnover: Why it matters and what we can do about it.* Learning Policy Institute. https://learningpolicyinstitute.org/product/teacher-turnover-report

Catts, H. W., Fey, M. E., Zhang, X., & Tomblin, J. B. (2001). Estimating the risk of future reading difficulties in kindergarten children: A research-based model and its clinical implementation. *Language, Speech, and Hearing Services in Schools*, *32*(1), 38–50.

Center for Education Policy Research. (2012). *Do low-performing students get placed with novice teachers?* Harvard University. https://hwpi.harvard.edu/files/sdp/files/sdp-spi-placement-memo.pdf

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. https://doi.org/10.1073/pnas.1315746110

Chambers, J. G. (1999). Measuring resources in education: From accounting to the resource cost model approach. U.S. Department of Education, National Center for Education Statistics. https://nces.ed.gov/pubs99/199916.pdf

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review*, *104*(9), 2633–2679. https:// doi.org/10.1257/aer.104.9.2633

Clark, M., Max, J., James-Burdumy, S., Robles, S., McCullough, M., Burkander, P., & Malick, S. (2022). *Study of teacher coaching based on classroom videos: Impacts on student achievement and teachers' practices* (NCEE 2022-006r). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/pubsearch/pubsinfo.asp?pubid=NCEE2022006

Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, *32*(6), 3–12. https://doi.org/10.3102/0013189X032006003

Collaborative for Academic, Social, and Emotional Learning (2023). *CASEL Program Guide.* https://pg.casel.org/

Cost Analysis Standards Project. (2021). *Standards for the economic evaluation of educational and social programs*. American Institutes for Research.

Cotton, K. (2000). *The schooling practices that matter most*. Association for Supervision and Curriculum Development; Northwest Regional Educational Lab. https://eric.ed.gov/?id=ED469234

Crosnoe, R., Kirpatrick-Johnson, M, & Elder, G.H. (2004). Intergenerational bonding in school: The behavioral and contextual correlates of student-teacher relationships. *Sociology of Education, 77*(1), 60–81. https://doi.org/10.1177/003804070407700103

Danielson, C. (1996). *Enhancing professional practice: A framework for teaching.* Association for Supervision and Curriculum Development.

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2016). *Effective teacher professional development*. Learning Policy Institute.

Deci, E. L., & Ryan, R. M. (2007). Active human nature: Self-determination theory and the promotion and maintenance of sport, exercise, and health. In M. S. Hagger, & N. L. D. Chatzisarantis, *Intrinsic motivation and self-determination in exercise and sport, human kinetics* (pp. 1–19). Champaign.

Delarosa, J. & Elias, J. (2022, December 6). *Forty-four percent of public schools operating without a full teaching staff in October, new NCES data show*. National Center for Education Statistics. https://nces.ed.gov/whatsnew/press_releases/12_6_2022.asp

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, *38*(3), 181–199. https://doi.org/10.3102/0013189X08331140

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.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Eccles, J. S., & Roeser, R. W. (2009). Schools, academic motivation, and stage-environment fit. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (3rd ed., Vol. 1, pp. 404–434). Wiley. https://doi.org/10.1002/9780470479193.adlpsy001013

Education Week. (2022, February 8). *Disrupted learning: The hard task of getting students back on track.* https://www.edweek.org/teaching-learning/disrupted-learning-the-hard-task-of-getting-students-back-on-track

Emmer, E. T., & Stough, L. M. (2001). Classroom management: A critical part of educational psychology, with implications for teacher education. *Educational Psychologist, 36*(2), 103–112. https://psycnet.apa.org/record/2001-17746-005

Evertson, C. M., Emmer, E. T., Sanford, J. P., & Clements, B. S. (1983). Improving classroom management: An experiment in elementary school classrooms. *The Elementary School Journal, 84*(2), 173–188. https://doi.org/10.1086/461354

Evertson, C. M., & Harris, A. H. (1999). Support for managing learning-centered classrooms: The classroom organization and management program. In H. J. Freiberg (Ed.), *Beyond behaviorism: Changing the classroom management paradigm* (pp. 59–74). Allyn and Bacon.

Foster, E. (2021). *Seeing teaching through a different lens.* Learning Forward. https://learningforward.org/wp-content/uploads/2021/01/mtpjan28.pdf

[redacted]

Gale, J., Alemdar, M., Cappelli, C., & Morris, D. (2021). A mixed methods study of self-efficacy, the sources of self-efficacy, and teaching experience. *Frontiers in Education*, *6.* https://doi.org/10.3389/feduc.2021.750599

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, *38*(4), 915–945. https://doi.org/10.3102/00028312038004915

[redacted]

Garrett, R., Citkowicz, M., & Williams, R. (2019). How responsive is a teacher's classroom practice to intervention? A meta-analysis of randomized field studies. *Review of Research in Education*, *43*(1), 106–137. https://doi.org/10.3102/0091732X19830634

Garrett, R., Zhang, Q., Citkowicz, M., & Burr, L. (2021). How Learning Forward's standards for professional learning are associated with teacher instruction and student achievement: A meta-analysis. Center on Great Teachers and Leaders at the American Institutes for Research.

George, S. V., Richardson, P. W., & Watt, H. M. (2018). Early career teachers' self-efficacy: A longitudinal study from Australia. *Australian Journal of Education*, *62*(2), 217–233. https://doi.org/10.1177/0004944118779601

Gershenson, S. (2016). Linking teacher quality, student attendance, and student achievement. *Education Finance and Policy*, *11*(2), 125–149. https://doi.org/10.1162/EDFP_a_00180

Gilman, R., & Anderman, E. M. (2006). The relationship between relative levels of motivation and intrapersonal, interpersonal, and academic functioning among older adolescents. *Journal of School Psychology, 44*(5), 375–391.

https://psycnet.apa.org/record/2006-12002-006

Greenberg, M. T., Weissberg, R. P., O'Brien, M. U., Zins, J. E., Fredericks, L., Resnik, H., & Elias, M. J. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologist*, *58*(6–7), 466–474. https://doi.org/10.1037/0003-066X.58.6-7.466

Gregory, A., Allen, J. P., Mikami, A. Y., Hafen, C. A., & Pianta, R. C. (2013a). Effects of a professional development program on behavioral engagement of students in middle and high school. *Psychology in Schools, 51*(2), 143–163. https://doi.org/10.1002/pits.21741

Gregory, A., Allen, J. P., Mikami, A. Y., Hafen, C. A., & Pianta, R. C. (2013b). *The promise of a teacher professional development program in reducing the racial disparity in classroom exclusionary discipline*. Center for Civil Rights Remedies National Conference. Closing the School to Research Gap: Research to Remedies Conference. Washington, DC. https://www.civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/state-reports/the-promise-of-a-teacher-professional-development-program-in-reducing-the-racial-disparity-in-classroom-exclusionary-discipline/gregory-teacher-development-ccrr-2013.pdf

Gregory, A., Allen, J. P., Mikami, A. Y., Hafen, C. A., Pianta, R. C. (2015) The promise of a teacher professional development program in reducing racial disparity in classroom exclusionary discipline. In Daniel J. Losen (Ed.), *Closing the school discipline gap: Equitable remedies for excessive exclusion* (pp 166–179). Teachers College Press.

Grissom, J. A., Egalite, A. J., & Lindsay, C. A. (2021, February). *How principals affect students and schools: A systematic synthesis of two decades of research.* The Wallace Foundation. https://wallacefoundation.org/sites/default/files/2023-09/How-Principals-Affect-Students-and-Schools.pdf

[redacted]

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.

Hill, H. C., & Papay, J. P. (2022). *Building better PL: How to strengthen teacher learning.* Research Partnership for Professional Learning. https://annenberg.brown.edu/sites/default/files/rppl-building-better-pl.pdf

Jacob, R., Zhu, P., & Bloom, S. H. (2010). New empirical evidence for the design of group randomized trials in education. *Journal of Research on Educational Effectiveness*, *3*(2), 157–198. https://doi.org/10.1080/19345741003592428

Kini, T., & Podolsky, A. (2016). Does teaching experience increase teacher effectiveness? A review of the research. Learning Policy Institute.

Klein, A. (2021, May 18). How to fix 7 fatal flaws in tech professional development. *Education Week*. https://www.edweek.org/leadership/how-to-fix-7-fatal-flaws-in-tech-professional-development/2021/05

Knight, J. (2017). The impact cycle: What instructional coaches should do to foster powerful improvements in teaching. Crown.

Kraft, M. A. (2019). Teacher effects on complex cognitive skills and social-emotional competencies. *Journal of Human Resources*, *54*(1), 1–36. https://doi.org/10.3368/jhr.54.1.0916.8265R3

Kraft, M. A., & Blazar, D. (2018). Taking teacher coaching to scale: Can personalized training become standard practice? *Education Next*, *18*(4), 68–74.

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*, *88*(4), 547–588. https://doi.org/10.3102/0034654318759268

Learning Forward. (2022). *Standards for professional learning.* https://standards.learningforward.org/standards-for-professional-learning/

Levin, H., McEwan, P., Belfield, C., Bowden, A. B., & Shand, R. (2018). *Economic evaluation in education: Cost effectiveness and benefit-cost analysis* (3rd ed.). Sage Publications.

Li, Y., & Lerner, R. M. (2011). Trajectories of school engagement during adolescence: Implications for grades, depression, delinquency, and substance use. *Developmental Psychology*, *47*(1), 233–247. https://doi.org/10.1037/a0021307

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). Wiley-Interscience.

Mangin, M. M., & Dunsmore, K. (2015). How the framing of instructional coaching as a lever for systemic or individual reform influences the enactment of coaching. *Educational Administration Quarterly, 51*(2), 179–213. https://doi.org/10.1177/0013161X14522814

National Center for Education Statistics. (2012). *Schools and staffing survey (SASS): 2011–12 SASS Questionnaires*. https://nces.ed.gov/surveys/sass/question1112.asp

National Center for Education Statistics. (2023a). Lower average mathematics score for eighth-graders than in 2019. U.S. Department of Education. https://www.nationsreportcard.gov/mathematics/nation/scores/?grade=4

National Center for Education Statistics. (2023b). Lower average reading score for eighth-graders than in 2019. U.S. Department of Education. https://www.nationsreportcard.gov/reading/nation/scores/?grade=4

National Research Council and Institute of Medicine. (2004). *Engaging schools: Fostering high school students' motivation to learn*. The National Academies Press. https://doi.org/10.17226/10421

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How Large Are Teacher Effects? *Educational Evaluation and Policy Analysis*, *26*(3), 237-257. https://doi.org/10.3102/01623737026003237

Pianta, R. C. (2016). Classroom processes and teacher-student interaction: Integrations with a developmental psychopathology perspective. In D. Cicchetti (Ed.), *Developmental psychopathology* (3rd ed., Vol. 4, pp. 770–814). Wiley.

[redacted]

Picucci, A. & Laughlin, T. (2019). Evidence-based coaching: Key drivers of scalable improvement district-wide. New Teacher Center.

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15*(3), 209–233.

Prothero, A. (2023, February 21). Student and teacher motivation, in charts. *Education Week*. https://www.edweek.org/leadership/student-and-teacher-motivation-in-charts/2023/02

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.

Reese, L, Jensen, B., & Ramirez, D. (2014). Emotionally Supportive Classroom Contexts for Young Latino Children in Rural California. *The Elementary School Journal*, *114* (4), 501–526. https://doi.org/10.1086/675636

Reeve, J. (2013). How students create motivationally supportive learning environments for themselves: The concept of agentic engagement. *Journal of Educational Psychology*, *105*(3), 579–595. doi:10.1037/a0032690

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*(2), 417–458. https://doi.org/10.1111/j.1468-0262.2005.00584.x

Roeser, R. W., Eccles, J. S., & Sameroff, A. J. (1998). Academic and emotional functioning in early adolescence: Longitudinal relations, patterns, and prediction by experience in middle school. *Development and Psychopathology, 10*(2), 321–352.

Roorda, D. L., Jak, S., Zee, M., Oort, F. J., & Koomen, H. M. (2017). Affective teacher–student relationships and students' engagement and achievement: A meta-analytic update and test of the mediating role of engagement. *School Psychology Review*, *46*(3), 239–261. https://doi.org/10.17105/SPR-2017-0035.V46-3

Samelo-Aro, K. (2017). Dark and bright sides of thriving—School burnout and engagement in the Finnish context. *European Journal of Developmental Psychology*, *14*(3), 337–349. https://doi.org/10.1080/17405629.2016.1207517

Sass, T. R., Hannaway, J., Xu, Z., Figlio, D. N., & Feng, L. (2010, November). Value added of teachers in high–poverty schools and lower–poverty schools. CALDER Working Paper No. 52. Center for Analysis of Longitudinal Data in Education Research. https://www.urban.org/sites/default/files/publication/33326/1001469-Value-Added-of-Teachers-in-High-Poverty-Schools-and-Lower-Poverty-Schools.PDF

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]. http://quantpsy.org/calc.htm

Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection. *Educational and Psychological Measurement*, *69*(3), 493–525. https://doi.org/10.1177/0013164408323233

[redacted]

Stuhlman, M., Mikami, A. Y., Hofkens, T., Allen, J., Pianta, R., & Smit, S. (2022). Integrating research-supported coaching practices into secondary teachers' team meetings: Early indications of potential to impact collaborations, classroom interactions, and student engagement. *Frontiers in Education*, *7*. https://doi.org/10.3389/feduc.2022.883226

Taie, S., & Lewis, L. (2023). Teacher attrition and mobility. Results from the 2021–22 Teacher Follow-Up Survey to the National Teacher and Principal Survey (NCES 2024-039). U.S. Department of Education. Washington, DC: National Center for Education Statistics. https://nces.ed.gov/pubs2024/2024039M.pdf

Taylor, B., Pearson, P., Peterson, D., & Rodriguez, M. (2003). Reading growth in high-poverty classrooms: The influence of teacher practices that encourage cognitive engagement in literacy learning. *Elementary School Journal, 104*(1)*,* 3–28.

Taylor, E. S., & Tyler, J. H. (2011). *The effect of evaluation on performance: Evidence from longitudinal student achievement data of mid-career teachers* (Working Paper 16877). National Bureau of Economic Research. https://www.nber.org/papers/w16877

Therriault, S. B., Walston, J., Li, Y., & Pan, J. (2020). *Relationships between schoolwide instructional observation scores and student academic achievement and growth in low-performing schools in Massachusetts* (REL 2020–026). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. https://ies.ed.gov/ncee/edlabs/regions/northeast/pdf/REL_2020026.pdf

Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing and elusive construct. *Teaching and Teacher Education, 17*, 783–805. https://doi.org/10.1016/S0742-051X(01)00036-1

U.S. Department of Education. (2014). *Fiscal year 2014 budget summary and background information*. https://www2.ed.gov/about/overview/budget/budget14/summary/14summary.pdf

Voisin, D. R., & Elsaesser, C. (2016). Brief report: The protective effects of school engagement for African American adolescent males. *Journal of Health Psychology*, *21*(4), 573–576. https://doi.org/10.1177/1359105314531607

[redacted]

[redacted]

Wells, K., & Foster, E. (2022, April). Strengths-based coaching supports teachers during COVID. *The Learning Professional*, *43*(2), 36–39. https://learningforward.org/wp-content/uploads/2022/04/strengths-based-coaching-supports-teachers-during-covid.pdf

Wenglinsky, H. (2000). *How teaching matters: Bringing the classroom back into discussions of teacher quality*. ETS Policy Information Center. https://www.ets.org/research/policy_research_reports/publications/report/2000/idxn

What Works Clearinghouse. (2022). *What Works Clearinghouse procedures and standards handbook, version 5.0.* U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE). https://ies.ed.gov/ncee/wwc/Docs/referenceresources/Final_WWC-HandbookVer5_0-0-508.pdf

Will, M. (2023, May 15). Teacher job satisfaction rebounds from last year's low. But there's still a ways to go. *Education Week*. https://www.edweek.org/teaching-learning/teacher-job-satisfaction-rebounds-from-last-years-low-but-theres-still-a-ways-to-go/2023/05

Wlodkowski, R. J., & Ginsberg, M. B. (1995a). *Diversity & motivation: Culturally responsive teaching.* Jossey-Bass. https://eric.ed.gov/?id=ED387036

Wlodkowski, R. J., & Ginsberg, M. B. (1995b). A framework for culturally responsive teaching. *Educational Leadership, 53*(1), 17–21. https://www.ascd.org/el/articles/a-framework-for-culturally-responsive-teaching

Wolfe, S. & Alexander, R. J. (2008, December). Argumentation and dialogic teaching: Alternative pedagogies for a changing world. Beyond Current Horizons.

Xia, J., Izumi, M. & Gao, X. (2015). School process and teacher job satisfaction at alternative schools: A multilevel study using SASS 2007–08 data. *Leadership and Policy in Schools, 14*(2), 167–203. https://doi.org/10.1080/15700763.2014.997935

Zins, J. E., Bloodworth, M. R., Weissberg, R. P., & Walberg, H.J. (2004). The scientific base linking social and emotional learning to school success. In J. E. Zins, R. P. Weissberg, M. C. Wang, & H. J. Walberg (Eds.), *Building academic success on social and emotional learning: What does the research say?* (pp. 3–22). Teachers College Press.