

Effective Instructional Coaching At-Scale for Middle Schools Using Jim Knight’s Coaching with Impact (CWI)

Project Narrative (Mid-Phase)

Contents

Introduction.....	1
A. Significance.....	2
B. Quality of the Project Design.....	6
B.1. Successfully Addressing the Needs of the Target Population.....	6
B.2. A Coherent and Sustained Program of Research and Development.....	9
B.3. Clearly Specified and Measurable Goals, Objectives, and Outcomes	11
B.4. Procedures to Maximize Efficiency and Productivity.....	14
C. Strategy to Scale.....	15
C.1 Strategies That Overcome Barriers to Scale	15
C.2. Broad Dissemination Mechanisms to Support Further Replication.....	17
D. Adequacy of Resources and Quality of the Management Plan.....	19
D.1. Exemplary Capacity to Bring the Project to Scale.....	19
D.2. Reasonability of Costs	23
D.3. Continuing the Work.....	23
D.4. Management Plan to Achieve the Objectives on Time and Within Budget.....	24
E. Quality of the Project Evaluation	25
E.1. Evidence That Meets WWC Standards Without Reservations	25
E.2. Clear Components, Mediators, Outcomes, and Measurable Thresholds.....	27
E.3. Valid and Reliable Performance Data on Relevant Outcomes.....	28
Bibliography	31

Introduction

When students return to their classrooms after COVID-19 closures and uneven virtual schooling, teachers will need to do their best teaching every day, informed by the most rigorous research the field has to offer. Especially in schools serving low-income students and students of color, effective teaching will be imperative. Even before the disruption, two thirds of American eighth graders from low-income families had failed to acquire basic proficiency in reading, writing, and mathematics, according to the National Assessment of Educational Progress (NAEP), and the picture is worse for students of color (National Center for Education Statistics [NCES], 2011, 2020a, 2020b). These poor outcomes have persisted for decades, even as the field’s knowledge has grown with regard to instructional practices that work. The aftershocks of the pandemic will intensify these challenges and deepen the need to connect research with practice.

For Absolute Priorities 1 and 3, the American Institutes for Research (AIR), Instructional Coaching Group (ICG), and Learning Forward (LF) propose a project to meet these challenges at scale through instructional coaching. With five district partners, the project will refine, scale, test, and sustain the Coaching With Impact (CWI) program, designed by ICG’s founder Jim Knight to help teachers select and successfully implement high-leverage, evidence-based practices.

ICG will train, monitor, and support the instructional coaches who serve the English language arts (ELA) teachers in middle schools in the partner districts. After a 1-year field test to refine its scaling mechanisms, ICG will begin a 3-year, at-scale implementation to help successive groups of ELA teachers achieve student goals for proficiency in writing (see Exhibit 1). In the first year, local coaches will partner with sixth-grade ELA teachers in a series of structured one-on-one coaching cycles driven by the teacher’s goals for students and practices from the CWI Instructional Playbook. After successfully completing two to three cycles with

each sixth-grade teacher, coaches will work with seventh-grade teachers in the second year and eighth-grade teachers in the third year.

AIR will lead the overall project. In addition to working with ICG, LF, and the district partners to recruit schools to implement CWI, AIR will carry out the evaluation activities. AIR’s iterative feedback to ICG will inform enhancements to ICG’s program tools and materials. During the 3-year at-scale implementation, AIR will evaluate the impact of CWI based on a randomized controlled trial with 80 schools in total, 40 of which will use CWI.

LF will lead the partners’ collaborative dissemination efforts, designed to bring key lessons and findings from the project to those who influence professional learning.

Exhibit 1. Number of Districts, Schools, and Teachers Implementing CWI

One-year field test to refine CWI	Three-year, at-scale implementation to evaluate the impact of CWI		
School year 2021–22	School years 2022–23, 2023–24, 2024–25		
3 districts, 6 schools	8 districts, 40 schools		
18 Grade 6 teachers	120 Grade 6 teachers	120 Grade 7 teachers	120 Grade 8 teachers

A. Significance

The significance of strategies to improve teacher effectiveness. Strategies to improve teacher effectiveness play an important role in efforts by policymakers and practitioners to boost student outcomes, as evident in federal, state, and local policies. These strategies can work in different contexts, adding value alongside curriculum or social-emotional learning (SEL) initiatives. Longitudinal research consistently shows that teacher effectiveness vary, within and across schools, in terms of improving student outcomes such as academic achievement, social and emotional competencies, and later-life outcomes (Chamberlain, 2013; Chetty, Friedman, & Rockoff, 2014; Gershenson, 2016; Kraft, 2019; 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 federal policy, the Every Student Succeeds Act calls on states and districts to ensure that such students are not taught disproportionately by less effective teachers.

But specific strategies that reliably improve teacher effectiveness are hard to find. For example, researchers have found that some teacher professional learning (PL) programs have an impact on student achievement, but many do not. For program reviews, see Kraft, Blazar, and Hogan (2018); Garet, Heppen, Walters, Smith, and Yang (2016); Garrett, Citkowitz, and Williams (2019); Gersten, Taylor, Keys, Rolhus, and Newman-Gonchar (2014); and Kennedy (2016). School systems invest in teacher PL every year, including billions of dollars from federal programs (U.S. Department of Education, 2014), so there is a clear need for strategies to improve teacher effectiveness that are proven, appealing, and scalable.

The potential to get much more out of the existing investment in instructional coaches. To improve teacher effectiveness, many districts invest in instructional coaches. Two-thirds (66%) of U.S. public schools have staff in coaching assignments, according to the latest nationally representative data (NCES, n.d.). Coaches are usually school-based and specialize in ELA or math. Districts typically train coaches on general techniques for coaching; it is uncommon for a district to train, monitor, and support coaches to use a specific approach.

There are many ways coaching can go astray (O’Keefe, 2017; Psencik, Mitrani, & Coleman, 2019; Simeral, 2018; Stricker & Culbertson, 2017; Wayne, 2019). If coaches struggle to engage teachers meaningfully, schools can easily assign coaches to other responsibilities.

Alarming, some districts now enumerate for schools what an instructional coach “is not” (e.g., “someone who teaches small groups or classes,” “a substitute teacher,” “a paraprofessional”).

The potential for coaching to have an impact is evident in the most rigorous studies—those that meet What Works Clearinghouse (WWC) standards—of coaching programs that share a core component: the use of detailed descriptions of proven instructional practice to guide teacher reflection on their teaching. There has been only one WWC intervention report for an instructional coaching program (WWC, 2015a). The program MyTeachingPartner–Secondary (MTP-S) demonstrated a positive impact on student achievement tests of 9 percentile points, across core subjects (Allen et al., 2011); a subsequent randomized controlled trial (RCT) produced the same impact (Allen et al., 2015). The coach draws on a description of proven instructional practice, the Classroom Assessment and Scoring System (CLASS) (Hamre & Pianta, 2005) tailored to middle and high school. Beyond this WWC intervention report, there are other individual studies of coaching that meet WWC standards and find positive effects for programs that guide teacher reflection based on descriptions of proven instructional practices. For example, Steinberg and Sartain (2015) use feedback based on Danielson’s Framework for Teaching (FfT) (Danielson, 1996). Garet et al. (2017) use both CLASS and FfT. In addition, WWC practice guides (e.g., WWC, 2015b) provide evidence supporting several discrete instructional practices (e.g., use of graphical aids, generating explanations, quizzing). These practices are particularly actionable for teachers, if coaches help them select, adapt, and implement the practices effectively in particular teaching contexts.

A challenge for instructional coaching programs: Scale. For a coaching program to benefit a large population, two things must happen: districts adopt the program; and coaches implement it with fidelity. District leaders are willing to fund instructional coaches but confront a marketplace saturated with models that claim to be “evidence-based.” Their decisions about

what to adopt depend a lot on brand names, marketing, and costs. The second facet of the scaling challenge, fidelity, is also fundamental (Kraft, Blazar, & Hogan, 2018). Most basically, program designers need to define what's expected of coaches, train them, and follow up with ongoing support that addresses the implementation challenges inherent in coaching (Wayne, 2019).

The significance of writing proficiency as a focal point for ELA coaching to close achievement gaps in middle school and beyond. While ELA achievement is a challenge for many middle school students, teachers struggle in particular with writing instruction—which is interwoven in the middle school ELA curriculum. According to the most recent NAEP results available, 34% of eighth grade students are proficient in reading (NCES, 2020) but only 27% are proficient in writing (NCES, 2012). Writing is seen as a key leverage point for the achievement gap in eighth grade, as writing proficiency rates for students from low-income families, Black students, and Latinx students are 12%, 11%, and 14%, respectively (NCES, 2012). Unlike NAEP reading scores, these low levels of proficiency have persisted since NAEP has been testing writing; there has been a 23–26 scale score point difference between White students and students of color for the past two decades (Salahu-Din, Persky, & Miller, 2008). These results are particularly significant given that the development of writing proficiency (a) has spillover effects on reading comprehension and fluency (Graham & Hebert, 2010; Graham & Santangelo, 2014; Zsigmond, 2015); (b) is used by teachers of other subjects to both acquire and assess student knowledge (Klein, Houg, & Bildfell, 2018); (c) is highly valued in the professional labor force (Pew Research Center, 2020); and (d) is a means to process and understand one's emotions and the emotions of others (Jones et al., 2017), a critical social-emotional life skill.

Summary and overview. This project will build knowledge and understanding about how to use instructional coaching effectively, at scale. The CWI program guides teacher

reflection using clear descriptions of evidence-based instructional practices. To ensure high-quality implementation at scale, it incorporates a system to train, monitor, and support instructional coaches. The project uses a version of CWI focused on student proficiency in writing, to provide a sensitive test of its impact on an outcome that plays a key role in achievement gaps and in preparing students to be informed, thoughtful, and productive citizens.

B. Quality of the Project Design

B.1. Successfully Addressing the Needs of the Target Population

The district partners (see Appendix I.1) chose middle school writing proficiency as the focus for this project, from among other versions of CWI (e.g., algebra, student behavior). Proficiency in writing at eighth grade is a strategic goal for them, requiring special attention in high-need middle schools. CWI will take place in middle schools with schoolwide Title I programs, where ELA teachers have access to a school-based instructional coach already. The coach will integrate CWI into their coaching with one grade level of ELA teachers each year, progressing from Grade 6 to Grade 8 across the 3-year, at-scale implementation (as detailed in Exhibit 1). During a year of CWI coaching, a teacher engages with the coach one-on-one in up to three coaching cycles, each following the CWI model and lasting 5–10 weeks, including weekly meetings and other interactions. The features of the CWI program are designed to successfully address the needs of the districts, teachers, and students (as discussed here) as well as coaches (see Section C.1).

District need for coach capacity. Despite curriculum and other initiatives in their high-need middle schools, the participating districts face persistently low ELA outcomes. The districts are eager to more effectively leverage their existing instructional coaches, who have received occasional summer training and limited ongoing support. To meet the need to build coach capacity, ICG trains, monitors, and supports each coach for 3 years (as detailed in section C.1).

The coach partners with a manageable group of teachers each year (usually 2–4 teachers), deepening their skill and experience with CWI and evidence-based instructional practices before turning to the next group of teachers and then the next. Districts thus gain a cadre of highly effective coaches as well as teachers.

Teacher need for self-directed PL that supports successful implementation of strong instructional practices. CWI reflects over a decade of research and development (see section B.3) to help meet teacher needs through three *Core Elements*: (1) the Impact Cycle, (2) an instructional playbook, and (3) guiding principles for coaching. To evoke the intrinsic motivation needed for lasting change to practice (Foley, 2011; Knight, Hock, Skritic, Bradley, & Knight, 2018; Yin & Lee, 2011), the three-phase *Impact Cycle* (see Exhibit 2) gives teachers control over their own PL; such control is critical for adult learning (Merriam, 2001) and teacher PL in particular (U.S. Department of Education, 2020). The *instructional playbook* entries provide a basis for coaches to help teachers understand what to implement and how. To successfully enact a new practice, teachers need an accessible, explicit description of the practice plus implementation guidance; each playbook entry describes an evidence-based instructional practice and actionable markers for implementation in the form of a checklist (see Appendix I.2.A for a sample playbook entry). The checklist anchors the iterative exchanges between teachers and coaches throughout the cycle. Finally, CWI coaches are trained in a set of *coaching principles* that help them meet each teacher’s need for a trusted partner whose stance facilitates partnership, dialogue, shared thinking, and guided reflection (Knight, 2015, 2019). With trust, teachers are open to receiving the help they need to notice and interpret student responses to their instruction and make effective adjustments (Downer, Jamil, & Maier, in press; Edutopia, 2016).

Exhibit 2. Snapshot of the Three Phases of the Impact Cycle

Identify: The teacher examines one or more videos of their instruction, student work samples, and other evidence of student learning to get a clear picture of the current classroom reality. With the help of their coach, the teacher selects a goal for the cycle that is student-focused and emotionally compelling for the teacher. The coach helps the teacher select an instructional practice to achieve their goal, drawing mainly from their coach’s instructional playbook.

Learn: The teacher selects the method(s) the coach will use to help them learn the practice—for example, modeling, coteaching, or collaboratively analyzing video samples. Teachers are then guided to begin to use the playbook checklist to implement the practice in their classroom and collect data on student responses (e.g., written performance tasks, exit tickets, and time-on-task counts).

Improve: The coach asks questions to help the teacher reflect on the data to see how the implementation of the practice seems to be shaping student engagement or learning and the extent to which their goals are being met. As a result of their analyses, teachers make informed adjustments to the practice in ways that improve outcomes.

With a focus on student writing proficiency, the CWI coach will help the teacher set a clear and measurable goal (e.g., ensure 90% of my students have the knowledge, attention, and persistence to achieve a score of 8 or more on the district informational writing rubric), select an evidence-based instructional practice from the playbook they believe can help them to address that goal (e.g., use a Model-Practice-Reflect cycle, organize classroom space to reduce distractions, praise effort), and learn to implement that practice effectively with coach partnership and evidence of student progress (e.g., time-on-task measures, writing samples).

Students’ needs for teaching that uses evidence-based instructional strategies responsively. To learn and develop, students need effective, engaging instruction. That is especially true in middle school, when engagement stagnates (Eccles & Roeser, 2009). For writing in particular, students need teaching that fosters self-esteem, perseverance, and the experience of success, because writing taps attention control, inhibitory control, planning, and perseverance (Graham & Harris, 2019). The practices in the instructional playbook meet students’ needs, helping the teacher organize instruction to engage students emotionally and cognitively. The specific practices come from authoritative evidence reviews, including the WWC (2016) practice guide on teaching secondary students to write effectively, additional

WWC practice guides, and consensus statements from cognitive science (e.g., Dunlosky et al., 2013) (see Exhibit 3). To ensure success with high-need students in particular, teachers have to implement these practices in a culturally responsive manner (Krasnoff, 2016). For that reason, the coaches in each district are supported to annotate the playbook to fit local student needs, as well as district-specific instructional frameworks and curricular resources.

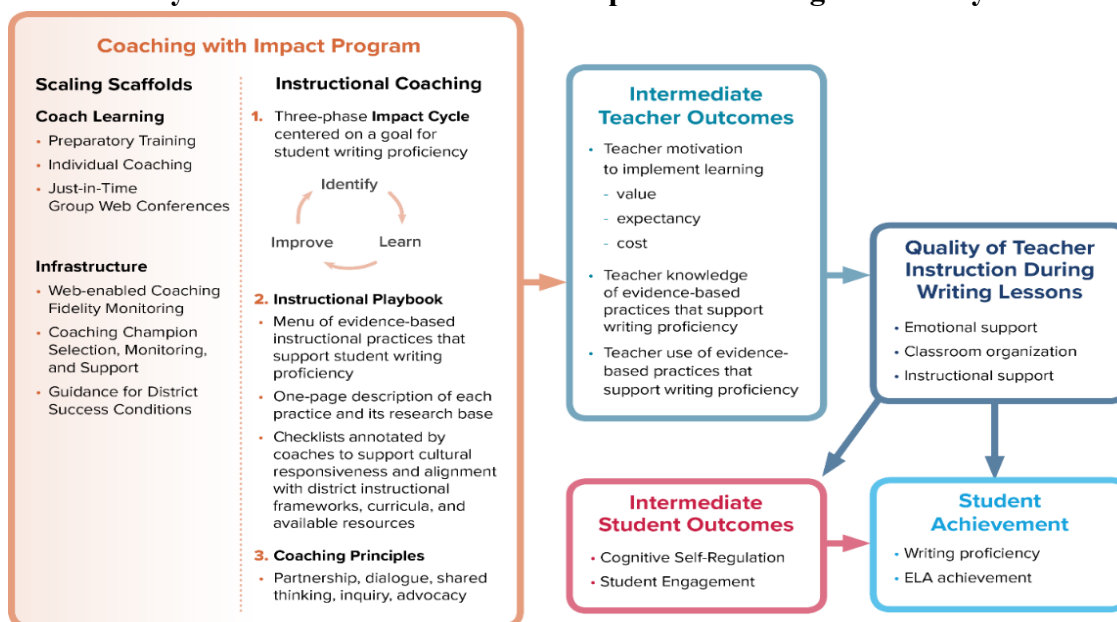
Exhibit 3. Sample Evidence-Based Instructional Practices in the CWI Instructional Playbook That Support Student Writing Proficiency

- Use a model-practice-reflect cycle (WWC, 2016)
- Integrate writing and reading to emphasize key writing features (WWC, 2016)
- Organize classroom time and tasks for distributed practice (Dunlosky, 2013; WWC, 2007)
- Organize classroom space to reduce distractions (APA, 2015)
- Provide explicit vocabulary instruction (WWC, 2008)
- Ask deep explanatory questions (Dunlosky, 2013; WWC, 2007)
- Provide specific, immediate statements of praise for desired behavior from students (APA, 2015)

B.2. A Coherent and Sustained Program of Research and Development

Underlying the success of the CWI program is a coherent theory of action (TOA), shown in Exhibit 4. Through a set of Scaling Scaffolds (elaborated in section C.1), the ICG trains, monitors, and supports coaches to provide instructional coaching that uses CWI’s well-specified process, playbook, and principles (see section B.1). The instructional coaching will motivate and support teachers to learn the practices from the playbook and implement them successfully. The use of the instructional practices, in turn, yields a measurably higher *Quality of Instruction* during writing lessons, which produces impacts on *Intermediate Student Outcomes* and *Student Achievement*. The project will use this initial theory as its guide, through the 1-year field test and the 3-year at-scale implementation, to carry out a sustained program of research and development that leverages regular feedback from AIR and the impact evaluation. The project will add to at least three ongoing lines of inquiry in the field.

Exhibit 4. Theory of Action for Evaluation of Impact on Writing Proficiency



How to scale instructional coaching with fidelity. One line of inquiry is how to scale instructional coaching with fidelity. In CWI, accomplished instructional leaders from ICG act as *coaching champions* who partner with coaches to monitor their progress and support ongoing coach learning to ensure quality implementation at scale. These program elements (i.e., the Scaling Scaffolds) reflect lessons from an investigation by ICG that monitored coaches and the quality of dialogue between coaches and teachers through weekly recordings (Knight et al., 2018). During the project, ICG will refine the design of the Scaling Scaffolds based on new implementation experiences and independent feedback from AIR. AIR will provide four cycles of feedback during the 1-year field test and two per year during the 3-year implementation.

How coaching can motivate teacher PL. A second line of inquiry that the project will continue is how to motivate teacher PL. Knight has written extensively on this question (see résumé in Appendix B). The project will explicitly measure teacher motivation to implement what they learn from CWI, using measures drawn from academic work to apply expectancy-value theory to teacher PL (Osman & Warner, 2020). The three core elements of the CWI

program (the impact cycles, instructional playbook, and coaching principles) are expected to ensure that (a) teachers perceive an emotionally compelling reason to make changes to their instruction (value), (b) teachers perceive that they *can* make effective changes with the support of playbook checklists and principled coaching (expectancy), and (c) teachers perceive that the benefits of making such changes outweigh the drawbacks (costs) as the playbooks and their coach help them understand the research behind the practices. See Appendix I.2.B for a description of how the Identify phase, for example, is designed to help teachers see the value of their effort in CWI in part through the use of classroom videos.

How to help teachers implement evidence-based practices effectively. A third line of inquiry ties into ongoing work on how to help teachers implement evidence-based instructional practices through coaching. For example, AIR is conducting two replication trials of MTP-S and a study of coaching based on mixed-reality simulation (Garrett, Smith, Yisak, & Griffin, 2020). AIR will use instruments from those studies, new instruments to measure use of the evidence-based practices that support student writing proficiency, and the CLASS-S measures of the quality of instruction (i.e., emotional support, classroom organization, and instructional support, as depicted in Exhibit 4). ICG will learn more about how its tools—especially the instructional playbook—can be refined to help teachers use proven practices effectively.

B.3. Clearly Specified and Measurable Goals, Objectives, and Outcomes

To advance these lines of inquiry, the project’s overall goal is to refine, scale, test, and sustain the CWI program and inform further work in the field. As shown in Exhibit 5, the management plan contains specific strategies, outcomes, and measures for each of its three key objectives.

Objective 1: Implement and refine CWI Scaling Scaffolds through a yearlong field test (2021–2022). An independently monitored field test of the CWI program, targeting sixth-

grade teachers for 1 year, will use regular objective feedback from AIR to provide ICG with actionable data to improve the Scaling Scaffolds, helping ensure strong implementation for the impact study and future use of CWI.

Objective 2: Implement CWI at scale over 3 years and determine its impacts (2022–2025). The implementation of CWI across 3 years will reach all ELA teachers in the 40 treatment schools in an 80-school RCT to assess the impact of CWI at scale. ICG aims to achieve a minimum effect size of 0.20 on student writing proficiency. A meta-analysis of PL programs that include instructional coaching suggests an average of 0.18 is reasonable (Kraft et al., 2018).

Objective 3: Sustain the project’s impact through broad dissemination and groundwork for future replication and scaling (2022 and beyond). The knowledge and insight gained through Objectives 1 and 2 will fuel dissemination that helps others apply the program in new contexts and subjects, incorporate key practices into other programs, and continue the project’s lines of inquiry.

Exhibit 5. Strategies, Outcomes, and Measures for Key Project Objectives

Strategies	Outcomes	Measures
Objective 1: Implement and refine CWI Scaling Scaffolds through a yearlong field test (2021–2022).		
Strategy 1.1. Identify 6 participating schools in partner districts. Confirm willingness to participate in field test.	School leaders agree to partner on the implementation of CWI and participate in data collection for 1 year.	Measure 1.1. District and school signatures on project memo of understanding (MOU) for 6 schools.
Strategy 1.2. Implement CWI with high degree of fidelity in 6 schools (2021–22 school year).	Each coach completes 2–3 Impact Cycles with 2–4 teachers in each school.	Measure 1.2. Online coaching logs and coach and teacher interviews indicate completed cycles.
Strategy 1.3. Collect implementation fidelity data and feedback as well as cost information (e.g., labor hours).	Comprehensive data on fidelity of implementation and costs of CWI are collected.	Measure 1.3. Teacher survey data, coach and coaching champion interview and log data, Impact Cycle fidelity scoresheet.
Strategy 1.4. Analyze and regularly share and discuss feedback with project partners.	Fidelity and cost data are summarized and shared during quarterly collaboration sessions.	Measure 1.4. Notes and artifacts (e.g., data visualizations) from 4 sessions.
Strategy 1.5. Refine Scaling Scaffolds based on feedback.	Training materials, web-based platform, and monitoring processes are revised.	Measure 1.5. Biannual memo summarizing revisions to each of 6 types of Scaling Scaffolds.
Objective 2: Implement CWI at scale for 3 years and determine its impacts (2022–2025).		

Strategies	Outcomes	Measures
Strategy 2.1. Identify participating schools in partner districts and additional districts if needed.	Principals at schools identified by partner districts support the project and agree to be randomly assigned.	Measure 2.1. District and school signatures on project MOU for 80 schools.
Strategy 2.2. Randomly assign schools to treatment and control conditions (40 schools per group).	Samples of treatment and control schools with baseline equivalence in key student, teacher, and school characteristics.	Measure 2.2. Number of schools in each group documented in random assignment memo.
Strategy 2.3. Implement CWI program and scaling supports in treatment schools (school years 2022–25).	CWI is implemented with high degree of fidelity and quality in 40 treatment schools.	Measure 2.3. All fidelity indicators in fidelity matrix meet adequate thresholds of fidelity for each implementation year.
Strategy 2.4. Measure and analyze fidelity of implementation including associated costs to implement.	Data on fidelity of implementation and labor and other direct costs are collected and analyzed.	Measures 2.4. All planned data sources are used to complete fidelity matrix and resource cost model.
Strategy 2.5. Analyze and regularly share and discuss feedback including preliminary outcomes data with project partners.	Biannual collaborative analysis/interpretation sessions are held and revisions to program and scaling materials are made based on data.	Measure 2.5. Notes and artifacts from 6 sessions.
Strategy 2.6. Measure and analyze treatment-control contrast in teachers' PL experiences.	Data on frequency and content of teachers' PL experiences collected and analyzed.	Measure 2.6. Teacher survey data indicates intended contrast in coaching frequency and content.
Strategy 2.7. Assess the impact of CWI on teacher and student outcomes.	Data on outcomes collected and analyzed as planned.	Measure 2.7. Data collection update; impact memo; impact findings meet WWC w/o res.
Objective 3: Sustain the project's impact through broad dissemination and groundwork for future replication and scaling (2022 and beyond).		
Strategy 3.1. Create research products (including written and video products).	Impact study interim and final reports; report briefs; cost analysis report; and journal, blog, podcast, videos, and newsletter articles are created based on project learnings.	Measure 3.1. Number of written and video products and number of times each product is cited or shared.
Strategy 3.2. Leverage distribution channels including social media outlets to share research products.	Partners collaborate to engage their internal and external networks to share research products.	Measure 3.2. Number of times each product is shared or cited.
Strategy 3.3. Present findings at research conferences and large practitioner convenings, including hosting special sessions with current and potential CWI users.	Researchers and participants share project learnings at 1 to 2 conferences per year beginning in Year 2, building knowledge among potential CWI users and the wider field.	Measure 3.3. Number of presentations made per project year; number of special conference sessions held.
Strategy 3.4. Convene leaders and coaches from network of participating districts to expand the use of CWI in their districts as well as to other districts.	CWI coaches and school leaders share experiences that help other users and inform those considering implementing CWI, resulting in CWI being spread to other schools, subject areas, and grade levels.	Measure 3.4. Meeting attendance records; meeting materials; memo summarizing input from partners and lessons; number of additional implementers.
Strategy 3.5. Distribute a quarterly e-newsletter to participant network.	Keep all study participants and stakeholders connected and updated on project developments.	Measure 3.5. Number of newsletters distributed, shared, and opened.
Strategy 3.6. Seek additional funds to support a future implementation of CWI in control group schools.	Fully fund implementation of CWI in all interested control group schools.	Measure 3.6. Number of control group schools implementing CWI after the project ends.

B.4. Procedures to Maximize Efficiency and Productivity

The CWI program will increase the efficiency and productivity of instructional coaches. In a current, six-district study at AIR, instructional coaches report salaries that range from \$81,650 to \$174,100 per year. But in the absence of strong training, monitoring, and support for coaches, instructional coaches often end up doing work of lower-paid staff, as discussed in section A. The CWI program designed a set of Scaling Scaffolds (detailed in C.1) to ensure coaches spend their time coaching, using the approach Knight specified (e.g., the 3-phases cycle, with action steps for each phase; the playbook; etc.) to improve teacher productivity and student outcomes.

The CWI program and its scaling mechanisms also improve efficiency and productivity, relative to traditional procedures, by leveraging technology for: (a) integration of distance-learning into coach preparation, which spaces learning over time to be more effective than training in a single multi-day event, (b) the online platform that guides and organizes key activities and interactions between the teachers and coaches (e.g., video-sharing, prompted reflection, use of CWI materials), and (c) remote, ongoing monitoring and support for coaches, enabled by the artifacts and coach notes captured in the online platform. Telecommunications technology will also make AIR's data collections efficient.

Finally, the project will improve the CWI program's efficiency and productivity, as explained in section B.3, based on implementation monitoring by ICG as well as the independent data from AIR. To help ICG consider program costs when refining the program, AIR will provide cost analyses using the Resource Cost Model (RCM), which models the ingredients of services as actually provided by the intervention (Levin & McEwan, 2000; Levin et al., 2017).

C. Strategy to Scale

C.1 Strategies That Overcome Barriers to Scale

Studies show that coaches and teachers can implement the Knight model for instructional coaching with fidelity (Knight et al., 2018; Knight & Cornett, 2009), but carefully designed supports are needed to ensure fidelity on a large scale. Therefore, ICG incorporates several specific strategies called *Scaling Scaffolds* (see the leftmost box of Exhibit 4) into the CWI program that address potential barriers to scaling (such as variation in coaching capacity and limited protected time for coaching activities). These include the following coach learning activities and infrastructure supports.

Preparatory training in which coaches learn CWI processes and adapt the instructional playbook. Through a series of webinars, assigned readings, and short quizzes, over the course of three days, coaches become familiar with the Impact Cycle, the instructional playbook and how to use it (fully detailed in Knight et al., 2019), and CWI coaching principles. They then meet together in person for three days to (a) practice using student work and data in the Impact Cycle; (b) learn to navigate the CWI web-based platform to review and share videos, keep logs of coaching activities and decisions, and organize coaching artifacts and resources; and (c) generate and rehearse explanations for the instructional practices in the playbook and collaboratively interrogate potential annotations, thereby learning the practices deeply.

A web-based platform designed to promote coaching fidelity. Powered by Swivl, the CWI web-based platform not only supports coaches' work, but also allows coaching champions to support and monitor coaches. Using the platform, champions review and comment on (a) video-taped coaching conversations that coaches have with teachers, (b) coaching logs, and (c) coaches' and teachers' time-stamped comments on teachers' classroom videos. Coaching champions use

Coaching Conversation Review Sheets to assess the fidelity of the coaching conversations (see Appendix I.2.C) and monitor teacher and coach progress using the online meeting logs.

Regular one-on-one coaching of coaches. Coaching champions host weekly 1-hour monitoring and support videoconferences with each coach they support. The champion uses a monitoring checklist to review the coach's logs and videos of coaching conversations in the web-based platform prior to the call to guide coach reflection on their work. The frequency of the one-on-one coaching activities gradually declines as coaches demonstrate fidelity to the model. After 3 years of coaching with fidelity, coaches are eligible to receive ICG Certification.

A program of follow-up, group-support web conferences that target coaching skills and processes. After coaches have completed their initial training and begun coaching, they participate as a cohort in weekly group interactive videoconferences with their coaching champion, who provides just-in-time content and tools for each aspect of the Impact Cycle as it occurs. The coaching champion also facilitates group problem solving and shares checklists to help coaches manage the complexity of coaching (see Appendix I.2.D for examples).

Careful selection, training, support, and monitoring of the coaching champion to implement coach learning scaffolds. The coaching champions are carefully selected, trained, and rigorously supported by ICG leadership. They all earned ICG Certification through a rigorous application process including 3 years of successful instructional coaching. They engage in biweekly group videoconferences led by ICG staff to share coaching strategies and resources. Prior to each call, the ICG project director reviews meeting logs kept by the coaching champion to detect fidelity problems and respond.

Guidance for school district partners to provide the conditions for successful implementation and scaling. Written guidance and meetings between personnel from the

district's central office and ICG ensure shared purposes and expectations (e.g., for coach and teacher time commitments). The meetings also allow review and discussion of implementation and scaling successes and challenges.

Measuring scaling success. Finally, to determine whether the scaffolds overcome the barriers to scaling, AIR will monitor the fidelity of the coaching as CWI is implemented at scale. When 94% of coaching cycles meet the threshold for adequate implementation as described in section E.2., we will be confident that the barriers have been addressed.

C.2. Broad Dissemination Mechanisms to Support Further Replication

LF will lead a collaborative effort that leverages the dissemination mechanisms of each partner and other mechanisms, such as *Education Week* and research journals. To do so, LF will tap staff at LF, AIR, ICG, and the partner districts to produce dissemination products (e.g., video testimonials, blogs, articles) that provide multiple perspectives and integrate the voices of CWI participants. In tandem, each partner will help the authors gain access to its dissemination mechanisms, and the partners will collectively amplify the reach of each product through their social media capacities. The dissemination will be designed to give practitioners, policymakers, and researchers opportunities to (a) use the program in new contexts, (b) incorporate key practices into other programs, and (c) incorporate project lessons into the knowledge base about teaching and learning. To that end, the products will emphasize the CWI program and its key features, the experiences of participants, and findings from the evaluation.

Learning Forward dissemination mechanisms. Among K–12 professionals who influence PL, LF is widely trusted as a source of relevant information and a facilitator of meaningful dialogue. Its membership includes central office and building administrators, teachers, PL providers, and researchers with a special interest in PL. LF will engage their

members as well as K-12 state and regional affiliates and other education audiences, using a combination of mechanisms: its electronic disseminations, which reach a 32,000-plus list of subscribers weekly (*PD in the News*) and monthly (*Connect*); its quarterly newsletter (*Tools for Learning Schools*); its journal (*The Learning Professional*); blogs on its website; and its “white paper” series. Most importantly for deep engagement with those interested in the project and coaching, LF will also use its annual conference, summer institute, and affiliate network meetings to support dissemination. The annual conference alone typically draws 3,000 attendees. LF will feature CWI participants in at least one session per year that targets other potential users. Finally, LF will launch a sustained dissemination effort to the participating districts, to build momentum for replication into additional schools and subject area. To do so, LF will hold an in-person convening for district delegates to discuss experiences with CWI and expanding its use and distribute a quarterly project newsletter to sustain interest and engagement.

ICG dissemination mechanisms. ICG leads the field in supporting instructional coaching and shares its latest practices and insights through its client-funded projects as well as publicly available books and journal articles. ICG will support replication of CWI in new contexts by communicating directly with recent ICG clients, who increasingly demand scalable models for instructional coaching and support for evidence-based instructional practices. ICG also will write articles, present at industry conferences, and leverage its annual Teaching Learning Coaching Conference, which typically draws hundreds of participants. These efforts will support districts and schools seeking to incorporate promising practices into their local programs, PL providers that seek to strengthen the designs of their programs.

AIR dissemination mechanisms. The comparative advantage AIR brings to the dissemination effort is its reputation and reach among researchers as well as policymakers and

practitioners who seek to use the best evidence. To support replication and development through its own dissemination mechanisms, AIR will feature products on the AIR website, which includes a mix of reports, videos, infographics, essays, and commentary; the AIR website logs hundreds of thousands of visits monthly. AIR will also use the e-mail networks it maintains for those who use evidence to support high-quality teaching. For example, AIR will leverage the dissemination networks of its Center on Great Teachers and Leaders and Regional Educational Laboratories, each of which engage policymakers, practitioners, and researchers.

Other dissemination mechanisms. The partners will also support replication and development indirectly, through third-party dissemination mechanisms such as journals and conferences. AIR is particularly experienced at dissemination targeting researchers, including peer-reviewed academic journals. The partners will also develop products for education trade publications (e.g., *Education Week*, *Educational Leadership*) and conferences (e.g., ASCD, National Center for Teacher Effectiveness [NCTE], Association for Middle Level Education).

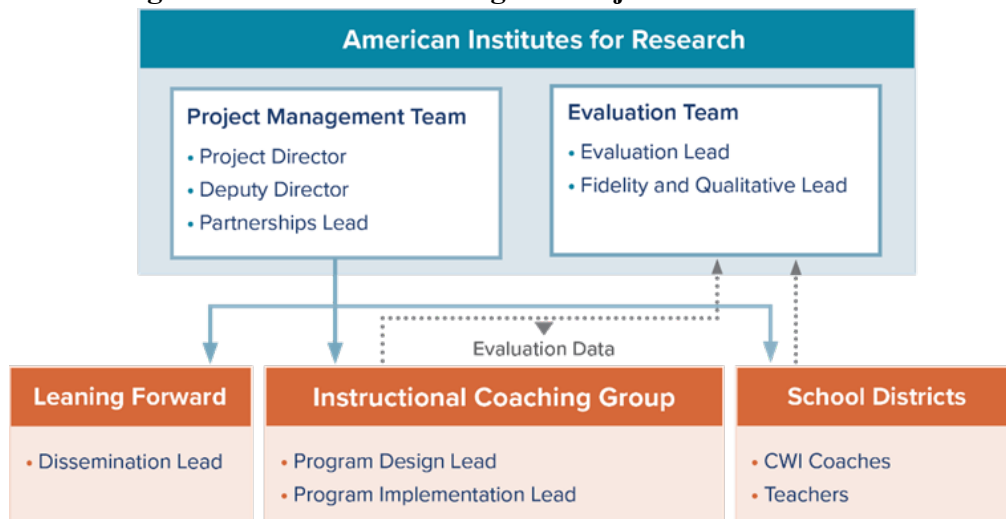
D. Adequacy of Resources and Quality of the Management Plan

D.1. Exemplary Capacity to Bring the Project to Scale

Each partner brings a track record that demonstrates the capacity to do its part to bring the project to scale, as well as experienced personnel organized as shown in Exhibit 6.

AIR is the lead organization for the project, responsible to the U.S. Department of Education for grant performance. AIR's role is to (a) oversee the subgrants to the ICG, the school district partners, and LF, ensuring coordination across the partners to achieve the project objectives; (b) recruit eligible schools from the school district partners; and (c) conduct the independent evaluation.

Exhibit 6. Management Structure to Bring the Project to Scale



Consistent with OESE’s guidance for independent evaluation (Abt Associates, May 2020), 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 implementation of the CWI program except evaluation feedback. This structure ensures the independence of key evaluation activities. In addition, AIR trains staff to report concerns about independence and tracks labor charges by task to ensure the division of labor between the project management tasks and the evaluation tasks.

AIR is uniquely qualified for this role. Using its extensive management infrastructure AIR has a distinguished track record of projects focused on teacher PL, including several led by AIR to scale up and evaluate PL programs. In each project shown in Exhibit 7, AIR coordinated across subcontracted organizations, including a PL provider and several districts; recruited participating districts, schools, and teachers; and conducted an independent evaluation to assess impacts. Across this work, AIR has developed tools and routines to provide independent, actionable feedback to those who design and deliver PL, including instructional coaching.

Exhibit 7. AIR-Led IES Contracts to Scale and Test Teacher PL Interventions (Completed)

Project name	Subcontractor(s) that provided the intervention	Number of districts and schools
The Impact of Providing Performance Feedback to Teachers and Principals	Danielson Group, Teachscape, University of Virginia, Discovery Education (coaching and performance feedback)	8 districts 127 schools
Middle School Mathematics Professional Development Impact Study	America's Choice and 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
Focusing on Mathematical Knowledge: The Impact of Content-Intensive Teacher PD	Intel Math (summer institute), Harvard University (video-based coaching)	6 districts 73 schools

AIR's management plan designates experienced staff from its teacher PL projects into lead roles, Leading its project management team, **Andrew Wayne, Project Director**, brings 20 years of experience studying teachers and teaching, including 15 leading projects that refine, scale, and test teacher PL programs. **Samantha Neiman, Deputy Project Director**, has project management certification and 13 years of experience that includes successful management of large-scale projects, including two current projects focused on instructional coaching. **Marlene Darwin, Partnerships Lead**, is an experienced ELA specialist and researcher with doctorate and master's degrees in ELA education. In AIR projects focused on teacher PL, she led district, school, and teacher recruitment and coordinated their activities with PL organizations.

Similarly, the AIR evaluation team leaders bring experience from AIR's work focused on teacher PL. **Mengli Song, Evaluation Lead**, has applied her research design and advanced quantitative methods expertise to these projects for 15 years. To complement her skills in impact evaluation, **Jane Cogshall, Fidelity and Qualitative Lead**, will apply her deep expertise in educator effectiveness and measuring implementation and providing feedback to PL providers.

ICG will provide all the necessary supervision, training, tools, and support that coaches need to implement CWI (summarized in section C.1). ICG is a global leader in instructional coaching, having worked with more than 100,000 coaches from around the world, delivering trainings and coaching of coaches. ***Jim Knight, Program Design Lead***, will lead the effort to refine CWI. He designed CWI and has refined materials based on his studies involving large implementations. In addition to being a national leader in the design of instructional coaching, demonstrated in articles and books, Jim Knight brings 5 years teaching writing at the college level. ***Sharon Thomas, Program Implementation Lead***, will lead the implementation activities and assist Jim Knight with refinements to the design. She is a National Board Certified English teacher and instructional coach. Thomas coordinates ICG's Coaching Certification process. She is also a Certified SIM Professional Developer in the area of writing. She founded the Cecil County [Maryland] Teacher Leadership Network and presents regularly on secondary school literacy and teacher leadership. ICG will work closely with Swivl, which uses its established electronic platforms for teacher PL to provide the secure digital environment for each phase of the coaching process.

LF will lead the dissemination and sustainability work, drawing on its leadership position and capacity to inform, reach, and engage those focused on K–12 PL (see learningforward.org). ***Tracy Crow, Dissemination Lead***, will direct this work, coordinating the dissemination effort across partners and participants as described in section C.2. As LF's Chief Strategy Officer, Tracy oversees the planning, creation, production, dissemination, and marketing of all LF products, including member newsletters, the journal *The Learning Professional*, and books.

With support from these organizations, the five district partners will (a) recruit and select suitable middle schools, (b) implement CWI, and (c) participate in all data collections (see Appendix I.3). To achieve the project's intended scale, each district brings a large number of

high-need middle schools (see Appendix I.1), and AIR will recruit additional or replacement districts as needed for the at-scale implementation that begins in 2022.

D.2. Reasonability of Costs

The project's cost control measures (see section B.4) and strong management toward clear objectives ensure that the project's cost is reasonable relative to its significant contribution to the field (see sections A and B.2). The additional benefits of the project accrue through students, teachers, and coaches.

Students gain proficiency in writing, which will benefit them significantly in later schooling and employment (Center on the Developing Child, n.d.). Given that writing skill can be a gateway to college (Common Core State Standards Initiative, 2010), 3 years of better instruction in writing can mean the difference between a high school and a college degree, which translates into almost \$1 million dollars of additional median lifetime earnings per student for a bachelor's degree and \$400,000 for an associate's degree (Carnevale, Rose, & Cheah, 2011).

The project provides similar benefits to future cohorts of students, through their teachers and coaches. The teachers will be more effective at teaching writing and ELA generally as they become confident users of proven instructional practices; taking into account all of the students they reach during the project period, the per-year cost of improved instruction for a student is \$91. The coaches will continue to guide teacher PL in their schools toward important student outcomes, leveraging all CWI-enhanced coaching skills, knowledge, and processes.

D.3. Continuing the Work

The project's benefits will continue into the future through the partners, who see the work as aligned to their mission. The ICG wants the CWI program and all its services to be effective, affordable, and widely used. It plans to use the project's final materials and procedures in future

implementations of CWI, after incorporating refinements based on the last year of implementation and evaluation findings. For the ICG, the project also fits with its ongoing efforts to refine its coaching model and apply it in powerful ways.

LF's mission includes helping states and districts identify scalable, effective PL focused on evidence-based practices. After the project, LF plans to continue convening users and potential users of CWI 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 PL that can support ongoing delivery of CWI. In addition, the cadre of coaches and teachers trained through the project in each school district partner 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 CWI, 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 CWI via the scaling strategy to the control teachers in each cohort after the project is over. AIR also plans to identify broader lessons from the evaluation to inform the development and enhancement of other PL programs. Finally, AIR will continue to pursue opportunities to build partnerships to scale, refine, and test teacher PL for high-need students, consistent with its mission to generate and use rigorous evidence that contributes to a better, more equitable world.

D.4. Management Plan to Achieve the Objectives on Time and Within Budget

The management plan ensures that the objectives will be achieved on time and within budget, because it assigns responsibility for each project objective *and* specific strategy (see section B.2) to a lead staff member who has a track record of success (see section D.1 for qualifications).

Appendix I.3 outlines the organization and person responsible, the timeline, and the milestones

for each strategy to complete the project objectives. The lead for each strategy will engage other partners and project personnel as needed to accomplish the strategy.

To implement the management plan, the AIR deputy project director (Ms. Neiman) will monitor progress and costs, ensure coordination across the partners, and adjust the plan as needed to assure quality and ensure success. To begin coordination efforts, the deputy project director will convene a project launch meeting upon award with all partners to clarify the management plan and lines of communication with school district partners. She will then set up regular task-based meetings with appropriate partners and team members and employ project management tools such as Airtable and Costpoint to monitor progress and costs.

E. Quality of the Project Evaluation

E.1. Evidence That Meets WWC Standards Without Reservations

AIR will conduct an independent evaluation to answer seven research questions (RQs) about the impact and implementation of the CWI program: (RQ1) What is the impact of CWI on teachers' motivation to implement PL and the quality of their practice during writing instruction? (RQ2) What is the impact of CWI on students' self-regulation, engagement, and achievement? (RQ3) To what extent is the impact of CWI on the quality of teacher practice moderated by teacher/classroom and school characteristics? (RQ4) To what extent is the impact of CWI on student achievement moderated by student, teacher/classroom, and school characteristics? (RQ5) To what extent is the impact of CWI on student achievement mediated by the quality of teacher practice? (RQ6) To what extent is CWI implemented with fidelity? (RQ7) What are the factors that hinder or facilitate the implementation of CWI? The 1-year field test (2021-22) will provide initial answers to RQ6 and RQ7 to inform refinement prior to the 3-year, at-scale implementation of CWI (2022–2025), which will answer all of the RQs.

AIR will use the 3-year, at-scale implementation to conduct a blocked cluster RCT, randomly assigning 80 middle schools to treatment and control conditions within each partner district for sufficient power (see Appendix I.4). All schools participating in the evaluation will be regular middle schools eligible for schoolwide Title I programs, which typically serve a large proportion of high-need students. Each school also will have at least two ELA teachers per grade and a staff member whose role includes instructional coaching for all ELA teachers.

For the 3 years following random assignment, the ELA teachers in both treatment and control schools will be subject to their districts' normal PL requirements and opportunities, but study teachers in treatment schools also will participate in CWI. The teachers will participate in three subgroups, based on the grade level of the majority of students they teach. Specifically, Teacher Cohort (TC) 1 will include study teachers primarily teaching Grade 6, who will participate in CWI in the first intervention year (IY1); TC2 will include teachers primarily teaching Grade 7, who will participate in CWI in IY2; and TC3 will include teachers primarily teaching Grade 8, who will participate in CWI in IY3. Thus, treatment teachers in the three cohorts will each receive a 1-year intervention in IY1, IY2, and IY3, respectively. For each cohort of teachers, teacher outcomes will be measured at the beginning of the intervention year—before CWI coaching starts (as baseline) and at the end of the intervention year (as outcome; see Appendix I.5). We plan to identify study teachers in all three cohorts before random assignment based on their expected teaching assignment during the intervention years and will not include any joiners in the teacher impact analyses. Thus, findings about CWI's impact on teacher outcomes based on the proposed school-level RCT, which is free of confounding factors, will have the potential to meet WWC standards without reservations. The evaluation sample also includes two student cohorts (SC1 and SC2). SC1 will include Grade 6 students taught by study teachers in IY1, followed longitudinally for 3 years—when they are in

Grades 6, 7, and 8, respectively (see Appendix I.5). Students in this cohort will be identified within the first 6 weeks of IY1 and will not include any late joiners. This will be our main student impact cohort, and impact findings based on this cohort have the potential to achieve the highest WWC rating—“meets WWC standards without reservations.” To explore teachers’ impacts on achievement in the year after they participate in CWI, we will examine a second student cohort: the Grade 6 students taught by study teachers in IY2. Findings based on this cohort can only meet WWC standards with reservations due to the potential risk of bias introduced by late joiners. This cohort will be followed longitudinally for at least 2 years and possibly 3 years if an optional follow-up year is included in the evaluation.

E.2. Clear Components, Mediators, Outcomes, and Measurable Thresholds

The evaluation design is informed by clearly articulated key components, mediators, and outcomes of CWI. As the TOA shows (see Exhibit 3), the key components of the CWI program include its Scaling Scaffolds and instructional coaching (coaching cycle, instructional playbook, and coaching principles). The TOA also specifies intermediate teacher outcomes (teacher motivation to implement PL, teacher knowledge of evidence-based practices that support writing proficiency, and teacher use of those practices) and the primary teacher outcome (quality of instruction during writing lessons). The intermediate teacher outcomes mediate CWI’s impact on the quality of instruction, which in turn mediates the program’s impact on intermediate student outcomes (i.e., cognitive self-regulation and student engagement) and subsequently student achievement outcomes (i.e., writing proficiency and ELA achievement).

The evaluation specifies measurable thresholds for acceptable implementation of the CWI program, which includes the Scaling Scaffolds and the instructional coaching. For the former, a CWI coach must complete 95% of the 6 days of preparatory training and participate in

at least three quarters of the just-in-time group coaching and individual coaching conversations. Program partners must provide 100% of the remaining Scaling Scaffolds described in section C.1. For the instructional coaching, coaches and teachers must complete at least two Impact Cycles per teacher during the year they are receiving the coaching with at least minimum fidelity. Fidelity levels will be determined using the Impact Cycle Fidelity Scoresheet (see Appendix I.6) scored by two raters on a randomly selected sample of coaching cycles. Cycles that score at the adequate level on at least eight of 10 indicators will be deemed to have met the fidelity criterion. These thresholds will be used to address RQ6 by assessing the fidelity of CWI implementation each year for each school and across the full sample and inform continuous improvement. In addition, the various types of implementation data collected for the proposed study (e.g., teacher surveys, coach interviews, and coaching logs) also will allow us to identify factors that hinder or facilitate the implementation and scaling of CWI (RQ7).

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

Teacher outcomes. According to the TOA underlying the CWI program (see Exhibit 4), the primary teacher outcome for CWI is the quality of teacher practice during writing instruction. For each teacher in the RCT sample, we plan to video-record—prior to coaching—one lesson in the early fall of the intervention year (as baseline) and two lessons in the spring of the intervention year. Given the intervention’s focus on writing instruction, we will work with teachers to schedule the video-recorded observations during writing instruction. The video-recorded lessons will be coded using CLASS-S, which is an instrument for measuring the quality of classroom interactions with rich evidence of reliability and validity (Bill & Melinda Gates Foundation, 2012; Goe, Bell, & Little, 2008). (See Appendix I.7 for the domains and dimensions of classroom practice measured by CLASS-S.). Certified CLASS-S 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. We will use the CLASS-S overall score as the primary teacher outcome measure and the three CLASS-S domain scores (Emotional Support, Classroom Organization, and Instructional Support) as supplemental teacher outcome measures. In addition, the evaluation will examine a key intermediate teacher outcome (teacher motivation to implement PL), drawing on a survey-based instrument with demonstrated validity and reliability (Osman & Warner, 2020). See Appendix I.8 for details about the three subscales associated with the measure: Expectancy for Success, Task Value, and Perceived Cost.¹

Student outcomes. As shown in the TOA (Exhibit 4), the CWI program is expected to affect both intermediate nonachievement student outcomes and achievement outcomes. As part of the project evaluation, we will examine two intermediate student outcomes: student-level cognitive self-regulation and classroom-level student engagement. We plan to measure students' cognitive self-regulation with two survey-based scales: the *Attention Control Scale* and the *Inhibitory Control Scale*. Both scales were developed for use with secondary school students for an internal AIR project and were shown to have sufficient reliabilities (Cronbach's alpha = 0.89 for both scales) and positive associations with students' SEL (see Appendix I.9 for details about the scales). We plan to measure student engagement with the CLASS-S dimension score for the Student Engagement dimension based on the classroom observation data described earlier.²

Given the intervention's focus on writing instruction, AIR will assess the impact of CWI on students' writing proficiency as the primary achievement outcome and on general ELA

¹ The TOA specifies two additional intermediate teacher outcomes: teacher knowledge and teacher use of evidence-based practices that support writing. We do not include them in the proposed teacher impact analyses because of the lack of reliable existing measures for these outcomes. We plan to develop reliable measures for these outcomes for during the project.

² To clearly distinguish measures of the teacher practice and student engagement, we will exclude the Student Engagement dimension when calculating the CLASS-S overall score as a measure of the quality of teacher practice.

achievement as the secondary achievement outcome. We will measure students' ELA achievement based on students' scores on state standardized ELA tests. Although state ELA tests typically include a writing component, states often use each student's performance (raw scores) on the writing component as one factor in estimating the overall ELA scale score for the student without generating a writing subscale score for each student. Therefore, we will administer a writing test to assess CWI's impact on students' writing proficiency. We have identified the ACT Aspire writing test as our measure of students' writing proficiency. As one of the five components of the ACT Aspire summative assessments, the ACT Aspire writing test is a vertically scaled, standards-based test that measures student growth and progress toward college and career readiness (see ACT, n.d.). It is available for Grades 3–10 and can be taken either online or via paper and pencil in 40- to 45-minute sessions. The *ACT Aspire Summative Technical Manual* (ACT, 2019) provides evidence demonstrating both the validity and reliability of the Aspire summative assessments, including the writing test.

The writing test and the student survey will be used to examine student outcomes after one year and three years of being taught by a teacher receiving CWI. These collections burden schools and are expensive so will focus on the main student cohort (SC1). At the beginning of IYI, AIR will randomly select a section taught by each teacher. The students from this section will be tested and surveyed at the beginning of IY1 (as baseline), the end of IYI, and then again at the end of IY3 (i.e., when the SC1 students are likely in different sections). Finally, AIR will collect extant ELA state test data in all years for both student cohorts, to explore possible impacts on broader, district-based measures of ELA achievement. (See Appendix I.10 for technical details on the impact analyses.)

Bibliography

Abt Associates. (2020, March 26). *EIR evaluation-related criteria & submission of draft EIR study design plans*. Bethesda, MD: Author.

ACT. (2019). *ACT Aspire summative technical manual* (Version 6). Iowa City, IA: Author.

Retrieved from <https://success.act.org/s/article/ACT-Aspire-Summative-Technical-Manual>

ACT. (n.d.). *ACT Aspire Summative Assessments*. Retrieved from

<https://www.act.org/content/act/en/products-and-services/act-aspire/about-act-aspire/summative-assessments.html>

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. Retrieved from <https://files.eric.ed.gov/fulltext/ED556046.pdf>

American Psychological Association, Coalition for Psychology in Schools and Education.

(2015). *Top 20 principles from psychology for preK–12 teaching and learning*. Retrieved from <http://www.apa.org/ed/schools/cpse/top-twenty-principles.pdf>

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

Retrieved from http://www.metproject.org/downloads/MET_Gathering_Feedback_Research_Paper.pdf

- Carnevale, A. P., Rose, S. J., & Cheah, B. (2011). *The college payoff: Education, occupations, lifetime earnings*. Washington, DC: The Georgetown University Center on Education and the Workforce. Retrieved from <https://1gyhoq479ufd3yna29x7ubjn-wpengine.netdna-ssl.com/wp-content/uploads/collegepayoff-completed.pdf>
- Center on the Developing Child. (n.d.). *InBrief: Executive function*. Retrieved from <https://developingchild.harvard.edu/resources/inbrief-executive-function/>
- 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.
- 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.
- Common Core State Standards Initiative. (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects. Appendix A: Research supporting key elements of the standards*. Retrieved from http://www.corestandards.org/assets/Appendix_A.pdf
- Cowan, J., Goldhaber, D., & Theobald, R. (2017). *Teacher equity gaps in Massachusetts* [ESE Policy Brief]. Malden, MA: Massachusetts Department of Elementary and Secondary Education.
- Danielson, C. (1996). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.

- Dong, N. and 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.
- Downer, J. T., Jamil, F., & Maier, M. (in press). Implications of information processing theory for professional development of early educators.
- 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. Steinber (Eds.), *Handbook of adolescent psychology* (3rd ed., pp. 404–434). Hoboken, NJ: Wiley.
- Edutopia. (2016). *Teachers gain from peer review*. Retrieved from <https://www.edutopia.org/practice/analyze-student-work-inform-instruction>
- Foley, L. S. (2011). Exploring K-3 teachers' implementation of comprehension strategy instruction (CSI) using expectancy-value theory. *Literacy Research and Instruction*, 50(3), 195–215.
- 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. 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., 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, 106–137.
- Garrett, R., Smith, T., Yisak, M., & Griffin, M. (2020). *Does practice make perfect? A randomized pilot study of a professional development program using mixed reality simulation to rehearse ambitious instruction*. Washington, DC: American Institutes for Research.
- 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*. Washington, DC: 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.
- Graham, S., & Harris, K. R. (2019). Evidence-based practices in writing. In S. Graham, C. A. MacArthur, & M. Hebert (Eds.), *Best practices in writing instruction* (3rd ed., pp. 3–28). New York, NY: Guilford Press.
- Graham, S., & Hebert, M. A. (2010). *Writing to read: Evidence for how writing can improve reading. A Carnegie Corporation Time to Act report*. Washington, DC: Alliance for Excellent Education.
- Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing*, 27(9), 1703–1743.
<https://doi.org/10.1007/s11145-014-9517-0>
- Hamre, B. K., & Pianta, R. C. (2005). Can Instructional and Emotional Support in the First-Grade Classroom Make a Difference for Children at Risk of School Failure? *Child Development*, 76, 949-967.
- 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.

- 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.
- Jones, S., Brush, K., Bailey, R., Brion-Meisels, G., McIntyre, J., Kahn, J., Nelson, B., & Stickle, L. (2017). *Navigating SEL from the Inside Out: Looking inside and across 25 leading SEL programs: A practice resource for schools and OST providers* (elementary school focus). New York: Wallace Foundation.
- Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, 86(4), 945–980.
- Klein, P. D., Houg, K. N., & Bildfell, A. (2018). Writing to learn (pp. 162–184). In S. Graham, C. A. MacArthur, & M. Herbert (Eds.), *Best practices in writing instruction*. The Guilford Press.
- Knight, D. S., Hock, M., Skritic, T. M., Bradley, B., & Knight, J. (2018). Evaluation of video-based instructional coaching for middle school teachers: Evidence from a multiple baseline study. *The Educational Forum*, 82, 1–18. Retrieved from <https://www.instructionalcoaching.com/wp-content/uploads/2019/08/Multiple-Baseline-Study.pdf>
- Knight, J. (2015). *Better conversations: Coach yourself to be more credible, caring, and connected*. Thousand Oaks, CA: Corwin Press.
- Knight, J. (2019). *The impact cycle: What instructional coaches should do to foster powerful improvements in teaching*. Thousand Oaks, CA: Corwin Press.

- Knight, J., & Cornett, J. (2009, April). *Studying the impact of instructional coaching*. Presentation delivered at the American Educational Research Association, San Diego, CA.
- Knight, J., Hoffman, A., Harris, M., & Thomas, S. (2019). *The instructional playbook: The missing link for translating research into practice*. Lawrence, KS: One Fine Bird Press.
- Kochhar, R. (2020). *Women make gains in the workplace amid a rising demand for skilled workers*. Washington, DC: Pew Research Center.
- Kraft, M. A. (2019). Teacher effects on complex cognitive skills and social-emotional competencies. *The Journal of Human Resources*, 54(1), 1–36.
- Kraft, M. A., & Blazar, D. (2018). Taking teacher coaching to scale: Can personalized training become standard practice? *EducationNext*, 18(4). Retrieved from <https://www.educationnext.org/taking-teacher-coaching-to-scale-can-personalized-training-become-standard-practice/>
- 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.
- Krasnoff, B. (2016). *Culturally responsive teaching: A guide to evidence-based practices for teaching all students equitably*. Portland, OR: Region X Equity Assistance Center Education Northwest. Retrieved from <https://educationnorthwest.org/sites/default/files/resources/culturally-responsive-teaching.pdf>
- Levin, H. M., & McEwan, P. J. (2000). *Cost-effectiveness analysis: Applications and methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Levin, H., McEwan, P., Belfield, C., Bowden, B., & Shand, R. (2017). *Economic evaluation in education: Cost-effectiveness and benefit-cost analysis*. Thousand Oaks, CA: Sage Publications.

Merriam, S. (2001). Andragogy and self-directed learning: Pillars of adult learning theory. *New Directions for Adult and Continuing Education*, 89, 3–14.
<https://eric.ed.gov/?id=EJ625870>

National Center for Education Statistics. (2008). *The Nation's Report Card: Writing 2007. National Assessment of Educational Progress at grades 8 and 12: National, state, and trial urban district results* (NCES 2008-468). Washington, DC: Institute of Education Sciences, U.S. Department of Education. Retrieved from
<https://nces.ed.gov/nationsreportcard/pdf/main2007/2008468.pdf>

National Center for Education Statistics. (2012). *The Nation's Report Card: Writing 2011. National Assessment of Educational Progress at grades 8 and 12* (NCES 2012-470). Washington, DC: Institute of Education Sciences, U.S. Department of Education. Retrieved from <https://nces.ed.gov/nationsreportcard/pdf/main2011/2012470.pdf>

National Center for Education Statistics. (2012). Figure: Percentages at or above each achievement level for writing, grade 8 by all students. *The Nation's Report Card: Writing 2011: Grade 8 National Results. National Assessment of Educational Progress at grades 8 and 12* (NCES 2012-470). Washington, DC: Institute of Education Sciences, U.S. Department of Education. Retrieved from
https://www.nationsreportcard.gov/writing_2011/g8_national.aspx?tab_id=tab2&subtab_id=Tab_1#chart

National Center for Education Statistics. (2012). Figure: percentages at or above each achievement level for writing, grade 8 by national school lunch program eligibility, 3 categories: 2011. *The Nation's Report Card: Writing 2011: Grade 8 National Results. National Assessment of Educational Progress at grades 8 and 12* (NCES 2012-470). Washington, DC: Institute of Education Sciences, U.S. Department of Education. Retrieved from

https://www.nationsreportcard.gov/writing_2011/g8_national.aspx?subtab_id=Tab_5&tab_id=tab2#chart

National Center for Education Statistics. (2012). Figure: The percentages of Asian eighth-graders performing at or above Proficient and at Advanced were higher than the percentages of students from the other racial/ethnic groups. *The Nation's Report Card: Writing 2011: Grade 8 National Results. National Assessment of Educational Progress at grades 8 and 12* (NCES 2012-470). Washington, DC: Institute of Education Sciences, U.S. Department of Education. Retrieved from

https://www.nationsreportcard.gov/writing_2011/g8_national.aspx?subtab_id=Tab_3&tab_id=tab2#chart

National Center for Education Statistics. (2020a). *NAEP report card: Reading. National achievement-level results* (Data compiled from National Assessment of Educational Progress, various years, 1992–2019 Reading Assessments). Washington, DC: Institute of Education Sciences, U.S. Department of Education. Retrieved from

<https://www.nationsreportcard.gov/reading/nation/achievement/?grade=8>

- National Center for Education Statistics. (2020b). *NAEP report card: Mathematics. National achievement-level results* (Data compiled from National Assessment of Educational Progress, various years, 1990–2019 Mathematics Assessments). Washington, DC: Institute of Education Sciences, U.S. Department of Education. Retrieved from <https://www.nationsreportcard.gov/mathematics/nation/achievement?grade=8>
- O’Keefe, B. (2017). *Primetime for coaching: Improving instructional coaching in early childhood education*. Washington, DC: Bellwether Education Partners. Retrieved from <https://files.eric.ed.gov/fulltext/ED585917.pdf>
- Osman, D. J., & Warner, J. R. (2020). Measuring teacher motivation: The missing link between professional development and practice. *Teaching and Teacher Education*, 92, 103064.
- Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods* 15(3), 209–233.
- Psencik, K., Mitrani, V., & Coleman, R. (2019). The power of coaching. *The Learning Professional*, 40(6), 54–57.
- 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.
- Salahu-Din, D., Persky, H., & Miller, J. (2008). *The Nation’s Report Card: Writing 2007* (NCES 2008–468). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

- 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.
- Simeral, A. (2018). 3 mistakes to avoid in your instructional coaching program. Retrieved from <https://www.k12insight.com/trusted/3-mistakes-avoid-instructional-coaching-program/>
- 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.
- Stricker, J., & Culbertson, J. (2017, June 15). *How to build a successful instructional coaching program* [Blog post]. Retrieved from <https://www.insighteducationgroup.com/blog/how-to-build-a-successful-instructional-coaching-program>
- TeachingWorks. (n.d.). *Certificate in practice-based approaches to methods instruction in teacher education* [Video file]. Retrieved from <http://www.teachingworks.org/pbtecertificate>
- U.S. Department of Education. (2020). *Proposed priorities, requirements, definition, and selection criteria*. Retrieved from <https://www.federalregister.gov/documents/2020/04/13/2020-07753/proposed-priorities-requirements-definition-and-selection-criteria-education-innovation-and>
- 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>

Wayne, A. J. (2019). *How to spread proven professional learning programs*. Retrieved from <https://learningforward.org/2019/08/how-to-spread-proven-professional-learning-programs/>

What Works Clearinghouse. (2007). *Organizing instruction and study to improve student learning* (NCER 2007-2004). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education. Retrieved from <https://ies.ed.gov/ncee/wwc/Docs/PracticeGuide/20072004.pdf>

What Works Clearinghouse. (2008). *Improving adolescent literacy: Effective classroom and intervention practices: A Practice Guide* (NCEE #2008-4027). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from <http://ies.ed.gov/ncee/wwc>.

What Works Clearinghouse. (2015a). *MyTeachingPartner–Secondary* (WWC Intervention Report). Washington, DC: Institute of Education Sciences, U.S. Department of Education. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/InterventionReports/wwc_mtp_061615.pdf

What Works Clearinghouse. (2015b). *Teaching strategies for improving algebra knowledge in middle and high school students* (NCEE 2014-4333). 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/wwc/Docs/PracticeGuide/WWC_Algebra_PG_Revised_02022018.pdf

- What Works Clearinghouse. (2016). *Teaching secondary students to write effectively* (NCEE 2017-4002). 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/wwc/Docs/PracticeGuide/508_WWCPG_SecondaryWriting_122719.pdf
- Yin, H.-B., & Lee, J. C.-K. (2011). Emotions matter: Teachers' feelings about their interactions with teacher trainers during curriculum reform. *Chinese Education and Society*, 44(4), 82–97. Retrieved from <https://doi.org/10.2753/CED1061-1932440405>
- Zsigmond, I. (2015). Writing strategies for fostering reading comprehension. *Procedia - Social and Behavioral Sciences*, 180, 1698–1703.