UC Irvine Writing Project's Pathway to Academic Success Program: An Education Innovation and Research (EIR) Expansion Grant Evaluation

EIR Expansion Grant #U411A180004

2024

Katrina R. Woodworth
Nicole L. Arshan
Haiwen Wang

SRI • 333 Ravenswood Avenue • Menlo Park, CA 94025-3493 • 650.859.2000 • www.sri.com

Acknowledgments:
This report was prepared by SRI International with funds provided by the UC Irvine Writing Project under an Education Innovation and Research (EIR) Expansion grant from the U.S. Department of Education (grant #U411A180004). However, these contents do not necessarily represent the policy of the U.S. Department of Education, and you should not assume endorsement by the U.S. Federal Government.

Suggested citation:
Woodworth, K. R., Arshan, N. L., & Wang, H. (2024). *UC Irvine Writing Project's Pathway to Academic Success Program: An Education Innovation and Research (EIR) Expansion Grant Evaluation*. SRI.

Contents

List of Exhibits

Abstract

Today's college and career-ready standards emphasize developing students' ability to interpret and analyze complex texts and to assert and defend claims relating to those texts in extended pieces of writing. By providing secondary English language arts teachers with professional development and instructional materials, the UC Irvine Writing Project's (UCIWP's) Pathway to Academic Success program (Pathway) seeks to support students to meet their state-adopted English language arts standards and graduate from high school prepared for college and work. In 2018, based on evidence of prior success in improving student writing, UCIWP received a federal Education Innovation and Research (EIR) Expansion grant to extend capacity to deliver Pathway teacher professional development in new contexts and validate the efficacy of the approach in these new settings. This technical report describes a school-level randomized controlled trial of Pathway, implemented by UCIWP and seven expansion Writing Project sites in seven states (AZ, CA, MN, OK, TX, UT, and WI). Independent SRI researchers randomly assigned 46 secondary schools to Pathway or a business-as-usual comparison group. In the Pathway schools, English teachers serving students in grades 7 through 11 participated in 1 or 2 years of professional development. SRI researchers measured student outcomes on the Analytic Writing Continuum for Literary Analysis (AWC-LA). Despite disruptions to program implementation during the COVID-19 pandemic, the study found positive effects on student writing quality, conventions, and productivity. The size, rigor, and independence of this study provide a strong evidence base to support Pathway's effectiveness in improving secondary students' academic writing at scale and in diverse contexts.

Introduction

After long being considered the neglected "R" of the traditional three "Rs" of education (National Commission on Writing, 2003), writing—in particular, nonfiction writing in response to texts—has become central to college- and career-ready standards, beginning with the Common Core State Standards in 2010. However, despite this shift in standards, writing remains a "neglected skill" in America's schools (Graham, 2019). The most recent national data on student writing achievement—now more than a decade old—showed that only about a quarter of U.S. secondary students were proficient in writing (National Center for Education Statistics, 2012). Although more recent national data on writing achievement does not exist, evidence suggests that overall achievement declined between 2019 and 2023 and that existing opportunity gaps have been exacerbated across subjects (Fahle et al., 2023).

The UC Irvine Writing Project's (UCIWP's) Pathway to Academic Success program (Pathway) has long history of effectively supporting writing instruction via carefully design professional development (PD) for secondary English language arts (ELA) and English language development (ELD) teachers (Kim et al., 2011; Olson & Land, 2007, 2008; Olson et al., 2012, 2017, 2020). The PD focuses on how to explicitly teach, model, and scaffold instruction in the cognitive strategies (or thinking tools) that research indicates experienced readers and writers access when they analyze and interpret complex texts. In addition to participating in PD, teachers are expected to implement at least three extended Pathway-developed lessons or "tutorials." By providing secondary ELA and ELD teachers with professional development and instructional materials, Pathway seeks to support students to meet their state-adopted ELA standards and graduate from high school prepared for college and work. In 2018, based on substantial evidence of prior success in improving student writing, UCIWP received a federal Education Innovation and Research (EIR) Expansion grant to extend capacity to deliver Pathway teacher professional development in new contexts and validate the efficacy of the approach in these new settings.

In this technical report, we describe a school-level randomized controlled trial of Pathway, implemented by UCIWP and seven expansion Writing Project sites in seven states (AZ, CA, MN, OK, TX, UT, and WI).[1] We begin with a brief review of relevant literature, followed by a detailed description of the impact study and an explanation of the extent which Pathway was implemented with fidelity. We then discuss the strategies UCIWP employed to build capacity and support implementation in new settings and the costs associated with program implementation. We conclude with a brief discussion of study implications.

[1] UCIWP and all expansion sites are affiliated with the National Writing Project, a network of 175 local Writing Project sites based in college and university campuses across the country. At the core of Writing Project work is developing and supporting teacher-leaders who in turn engage in professional development with their peers.

Note on the study context: UCIWP's expansion effort and SRI's evaluation of Pathway over three school years (2019/20 through 2021/22) were interrupted by the COVID-19 pandemic. Pandemic-related disruptions to Pathway implementation included school closures; virtual and sometimes asynchronous professional development and classroom instruction; unusually high teacher and student absenteeism; and very limited access to substitute teachers, which impacted teachers' release time to participate in professional development. To help the expansion sites navigate these disruptions, UCIWP program developers had to adapt Pathway. They supported sites to provide the professional development in shorter increments that were more suitable to a virtual environment. They also developed new, online instructional resources that teachers could implement with their students in a virtual learning environment.

Research Literature

We situate this study within three areas of prior research: writing instructional practices (including prior studies of Pathway), teacher professional development, and scaling successful interventions.

Writing Instructional Practices

The What Works Clearinghouse (WWC) practice guide for secondary writing instruction draws on 15 studies with high internal validity (Graham et al., 2019), three of which were studies of Pathway (Kim et al., 2011; Olson & Land, 2008; Olson et al., 2012, 2017). The practice guide authors offer three recommendations for writing instruction, with the Pathway studies contributing to each of the findings:

- Explicitly teach appropriate writing strategies using a Model-Practice-Reflect instructional cycle.
- Integrate writing and reading to emphasize key writing features.
- Use assessments of student writing to inform instruction and feedback.

A more recent synthesis of causal studies of writing instruction by Slavin et al. (2019) points out a weakness in the evidence base for these recommendations: While the supporting studies had high internal validity, most were small in scale, executed over a short period, or demonstrated results for only a subpopulation of students (e.g., students with learning disabilities). In their review, the authors extended the focus to writing instruction in grades 2–12 and modified the inclusion criteria to studies that extended for at least 12 weeks and involved at least 30 students and 2 teachers in each treatment condition, resulting in 14 studies of 12 writing programs, including Pathway (Kim et al., 2011; Olson et al., 2012, 2017). Despite the different grade levels and inclusion criteria, the Slavin et al. (2019) review yielded similar recommendations for writing instruction to those in the 2019 WWC practice guide, as well as two new recommendations: (1) use student-centered approaches, such as a focus on building student motivation, and (2) provide extensive PD in which teachers write as if they were students. Prior research on Pathway supports both recommendations.

Teacher Professional Development

A critical strategy for increasing the use of effective writing instructional practices is to provide effective PD experiences for teachers. Providers of PD have had limited success supporting teachers to improve instruction in ways that translate to increased student learning, even when the support has adhered to design features considered to be indicators of quality (e.g., Garet et al., 2016). These design features have included content focus, opportunities for active learning, coherence with instructional context, sustained duration, collective participation, coaching, and feedback for teachers (Desimone et al., 2002; Garet et al., 2001; Darling-Hammond et al., 2017).

More recently, the Research Partnership for Professional Learning (RPPL) summarized the evidence on features of effective teacher professional learning that distinguishes two domains: the "how" (format) and the "what" (content; Hill & Papay, 2022). For the how, RPPL cites research that suggests three professional learning formats can lead to improved instructional quality: "(1) built-in time for teacher-to-teacher collaboration around instructional improvement; (2) one-to-one coaching, where coaches work to observe and offer feedback on teachers' practice; and (3) follow-up meetings to address teachers' questions and fine-tune implementation" (p. 2). As described in our discussion of implementation fidelity and in keeping with key tenets of the National Writing Project, Pathway builds in substantial time for teacher collaboration around instruction and follow-up meetings to address teachers' questions. Pathway does not involve one-on-one coaching.

For the what, RPPL cites evidence that effective professional learning focuses on "(1) building subject-specific instructional practices rather than building content knowledge alone; (2) supporting teachers' instruction with concrete instructional materials like curricula or formative assessment items rather than focusing only on general principles; and (3) explicitly attending to teachers' relationships with students" (Hill & Papay, 2022, p. 2). Again, Pathway is well aligned, as the PD is designed to focus on specific writing instructional practices and with the expectation that teachers implement concrete instructional materials (i.e., tutorials) that are constructed to be of high interest to secondary students.

Finally, Kennedy's (2016) review of the research on teacher PD found that the "specific design features that are presumed to define high quality PD … may be unreliable predictors of program success" (p. 971). The author urges the field to attend to who provides the PD and teacher motivation to participate in the PD. In the case of the Pathway EIR Expansion grant, the PD offered by the expansion sites was typically provided by experienced teacher leaders with long affiliations with their local Writing Project sites, and teachers opted in to the PD, although some may have been strongly encouraged by their school or district leaders.

Scaling Interventions

EIR Expansion grants support interventions with a strong evidence base for one population or setting with the intention of implementing and testing the interventions at scale (i.e., at the national level). Demonstrating positive outcomes for large and diverse populations is often the primary purpose of scaling (McDonald et al., 2006; Morel et al., 2019). Despite the widely shared goal of reaching large numbers of students with evidence-based interventions, there is no formal definition of implementation at scale, and most evaluations of writing interventions that met WWC guidelines have had small samples that would fail to meet any reasonable definition of the term. Of the 14 studies included in the WWC practice guide for secondary writing (Graham et al., 2019), nine measured outcomes in six or fewer schools or, if school sample sizes were not included, for fewer than 100 students. By contrast, this Pathway impact study included

208 teachers in 46 schools in 8 districts across 7 states, and it builds on an earlier study that included 230 teachers across 7 Southern California districts (Woodworth et al., 2017).

In her influential framework on "rethinking scale," Coburn (2003) argues that expanding interventions to new settings is a "necessary but insufficient condition for scale" (p. 4). While Coburn refers to this dimension of scaling as "spread," she also identifies additional dimensions of scaling, including shifting reform ownership from the initial developers of an intervention to others, effecting a depth of change in teachers' beliefs and practices, and promoting the sustainability of intervention effects after initial implementation. As presented in the section describing the Scale-Up Evaluation, by developing the capacity of Writing Project sites to implement Pathway in their services areas, UCIWP attempted to help teacher-leaders develop a sense of ownership for the work (e.g., by first implementing Pathway tutorials in their own classrooms) and to encourage ongoing support for Pathway and related instructional practices.

Impact Study

Study Description

The evaluation is a 2-year school-randomized controlled trial (RCT) that investigates the impact of the Pathway Project on middle and high school students' writing. The confirmatory research question investigates the impact of the program on grade 7–11 students' overall writing quality on a text-based analytic writing task, wherein students read a piece of literary nonfiction and write an analysis. This writing is scored using the Analytic Writing Continuum for Literary Analysis (AWC-LA), which serves as the primary measure of student achievement. Exploratory questions include whether the impact on writing quality differs for different student subgroups, whether Pathway has an impact on writing conventions and productivity from the same analytic writing task, and whether Pathway has an impact on student state ELA/reading achievement test score for students in tested grades.

Research Questions

Our confirmatory research question was:

- What is the impact of Pathway on writing quality for students in grades 7–11 whose schools were randomly assigned to participate in Pathway compared to students in grades 7–11 whose schools were randomly assigned to business as usual?

We also posed the following exploratory research questions:

- Do student baseline characteristics (e.g., gender, socioeconomic status, eligibility for English learner services) moderate any estimated impacts of Pathway on student writing?
- What is the impact of Pathway on student state ELA achievement test scores for students in tested grades?
- Does the impact of Pathway on writing quality for students in grades 7–11 compared to business-as-usual instruction differ by timing or duration of intervention?
- What is the impact of Pathway on writing conventions for students in grades 7–11 whose schools were randomly assigned to participate in Pathway compared to students in grades 7–11 whose schools were randomly assigned to business as usual?
- What is the impact of Pathway on writing productivity for students in grades 7–11 whose schools were randomly assigned to participate in Pathway compared to students in grades 7–11 whose schools were randomly assigned to business as usual?

Research Design

This study is a 2-year, school-level cluster RCT. With guidance from the SRI research team, local Writing Project sites recruited pairs of schools serving the same grade levels and students with

similar achievement and demographic characteristics. In fall of the first year, prior to randomization, the local Writing Project sites recruited and consented teachers within those schools. Once the schools and teachers were confirmed, the SRI research team identified focal classes for those teachers. We identified the student sample either through student rosters or—when possible to conduct before randomization—participation in a baseline writing assessment (described below).

Once the student sample was established, SRI randomized the schools, inclusive of participating teachers and the students in their focal classes. SRI conducted randomization within the paired blocks in which schools had been recruited. This pair randomization was conducted to provide better equivalence across baseline indicators of key outcomes and contextual factors.[2] Additionally, because the local Writing Project sites planned to offer professional development to comparison group teachers upon the completion of the evaluation (i.e., delayed treatment), randomization within pairs ensured that each local Writing Project site would serve an equal number of schools in the treatment (early start) and comparison (delayed start) groups.

The local Writing Project sites and their partner districts launched the intervention over 3 years: fall 2019 (Cohort 1), fall 2020 (Cohort 2), and fall 2021 (Cohort 3), respectively.[3] For each cohort of schools, the SRI research team randomized schools, together with participating teachers and their students in two focal classes, as described above. In Year 2, for Cohort 1 and Cohort 2, these randomized teachers' classrooms contained both students from the original randomly assigned sample and in-movers. While Pathway teachers implemented the program for the second time, some students experienced it for the first time, and others continued in it for a second year. Unlike the Cohort 1 and 2 Pathway schools, Cohort 3 Pathway schools only received one year of the intervention (because of COVID-19 pandemic-related challenges).

Setting

UCIWP recruited seven additional local Writing Project sites that (1) served a high proportion of English learners and (2) represented a broad range of social, demographic, and policy contexts. The local Writing Project sites in turn recruited public school districts in their service areas, with the goal of serving districts with high proportions of English learners in grades 7–11 and a desire to implement the Pathway program. The sites and districts were:

- Arizona State University: Phoenix Union High School District
- University of California, Irvine: Long Beach Unified School District

[2] The research team had additionally hoped that randomizing within school-by-grade blocks would help the research design by allowing students to move from a focal classroom in Year 1 to a same-condition focal class in Year 2.
[3] The study was originally designed to have only two cohorts, with Cohort 1 (a single district, being served by the UCIWP program developer) being treated as a demonstration site and all other sites implementing after a year of capacity development and opportunity to learn from the program developers at UCIWP. As Cohort 2 launched during the COVID-19 pandemic, some local Writing Project sites encountered challenges with recruitment. As a result, two sites launched in fall 2021 as Cohort 3. However, because of the grant timeline, Cohort 3 treatment schools only received 1 year of Pathway professional development.

- University of California, Los Angeles: Los Angeles Unified School District
- University of Minnesota: Anoka-Hennepin School District
- Oklahoma State University: Oklahoma City Public Schools
- Texas State University-San Marcos: Judson Independent School District
- Brigham Young University: Nebo School District
- University of Wisconsin-Milwaukee: Milwaukee Public Schools

Intervention Condition

Program developers at UCIWP trained their peers at seven local Writing Project sites affiliated with the National Writing Project to develop the capacity to offer the Pathway program to teachers in their respective services areas. Pathway provides teachers with both professional development and curricular resources. Pathway was designed as a 2-year program. Because the second year was designed to reinforce the work of the first year, the core components of the program are largely the same. Because of pandemic-related challenges, two of the seven expansion sites implemented Pathway for just 1 year. Additionally, Pathway professional development was designed to be offered in-person over full days. Again, because of the pandemic, the professional development delivery was adjusted to accommodate remote learning. Specifically, the UCIWP Pathway developers (1) supported sites to provide the professional development in shorter increments that were more suitable to a virtual environment and (2) developed new online, student-facing instructional resources that could be implemented in a virtual learning environment.

Notwithstanding pandemic-related adjustments, four key features define the Pathway program:

- *Professional development content*. The program includes two introductory "tutorials" (or multiday lessons) on cognitive strategies, one focused on reading and one on writing; resources to support in-class book clubs; a revision tutorial; and explicit directions for teachers to lead their classes through a multiple-draft revision process.
- *Scaffolded approach*. The program is designed to support teachers' learning by modeling Pathway lessons for them during professional development events and to facilitate transfer to the classroom by providing ready-to-use instructional materials.
- *Formative feedback*. The program requires teachers to administer a pretest writing assessment at the beginning of each year and a posttest writing assessment at the end of each year. The program then provides teachers with feedback on their students' pretest writing. This program feature ensures students are set up to engage in the multiple-draft revision process and provides teachers with guidance on next steps for their instruction.
- *Classroom implementation*. Use of the tutorials to teach students to use the cognitive strategies and revise their writing is at the core of the program.

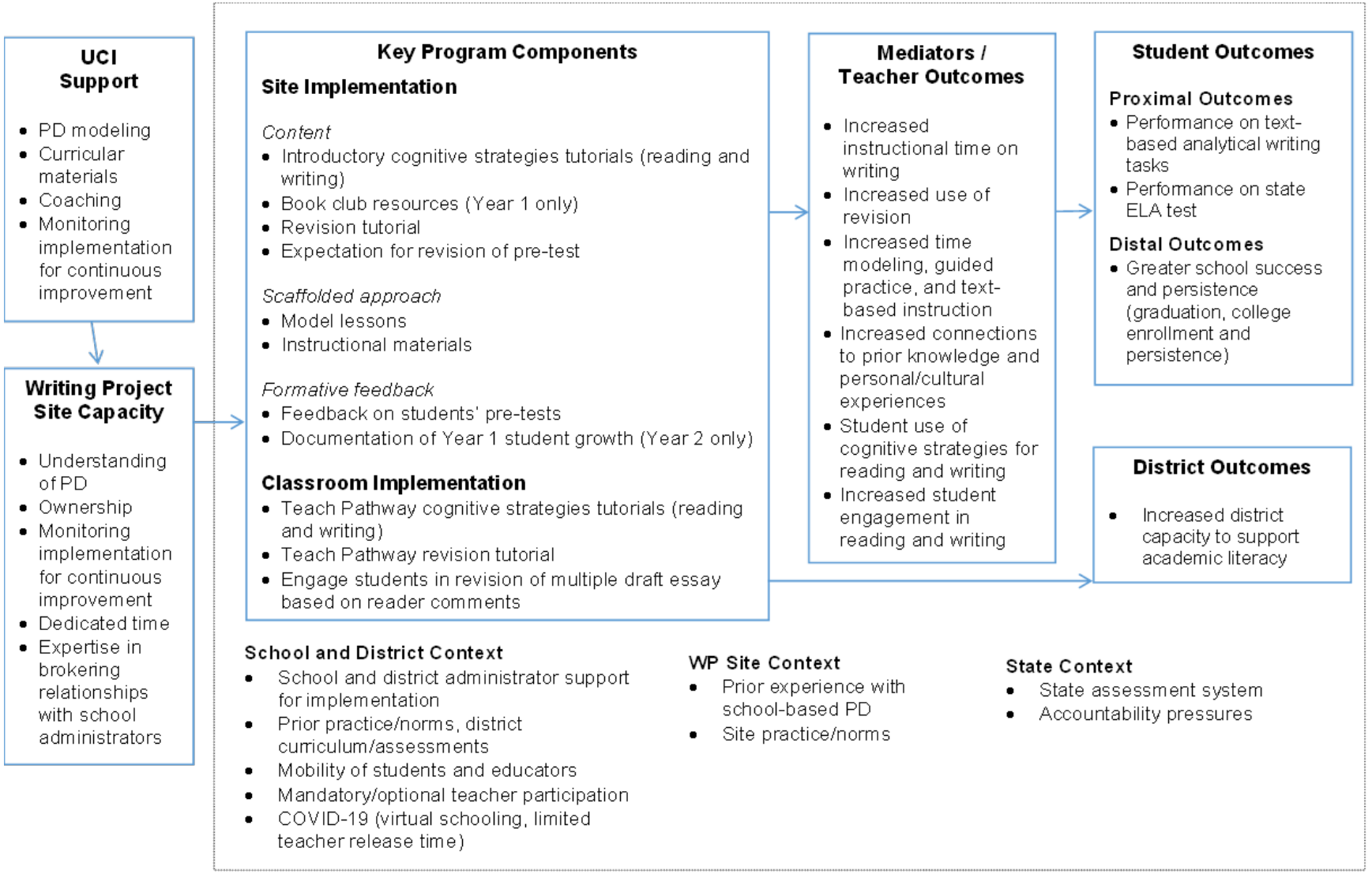
In addition, when the Pathway program extends into a second year, Year 2 includes the following additional program component:

- *Formative feedback*. In the summer between Years 1 and 2, Year 1 students' writing (pretests and posttests) is scored so that at the beginning of Year 2, teachers receive their students' scores from the prior year. The idea is that seeing growth from the pretest to the posttest demonstrates the efficacy of Pathway strategies and motivates teachers to implement the program with even greater fidelity in Year 2.

Specific information about implementation thresholds and measures of fidelity are discussed later in the Fidelity of Implementation Study section.

Pathway program developers and SRI researchers collaborated on a logic model that depicts the core program components as well as the inputs (e.g., UCIWP support for expansion sites, expansion sites' capacities), mediators (i.e., teachers' instructional practices), contextual factors that were expected to influence implementation (e.g., district curriculum and state assessments), both proximal outcomes (i.e., improved writing skills) and distal outcomes (e.g., greater school success and persistence), and system outcomes (e.g., increased district capacity to support academic literacy; Exhibit 1).

Exhibit 1. Logic Model for the Pathway to Academic Success Program



Comparison Condition

Comparison schools are those schools that were eligible for the intervention but were randomly assigned to the control condition. Teachers in these schools did not receive the Pathway professional development or associated materials during the 1 or 2 years of the study but were otherwise free to participate in professional development activities that were part of their professional responsibilities in their schools and districts (i.e., business-as-usual condition). During the study, comparison teachers received stipends for data collection activities. Upon completion of the study, the comparison schools were invited to become Pathway schools and comparison-school teachers were able to receive the professional development, materials, and support associated with the Pathway program.

Not all teachers and students within comparison (or treatment) schools were included in the RCT. Local Writing Project sites recruited as many eligible teachers as possible within participating schools for the program. SRI then randomized schools after (1) collecting student rosters for all classrooms in the school and (2) identifying participating teachers' focal classrooms (as described below).

Recruitment and Random Assignment

Within their partner districts and in consultation with SRI staff, **local Writing Project sites recruited pairs of schools** (e.g., 2 middle schools, 2 high schools) with similar achievement and demographics. SRI staff considered factors such as total enrollment, English learner enrollment, grade levels served, school schedules (e.g., half-year blocks), overall performance, and academic themes when pairing schools. After identifying study teachers, focal classes, and students (described below), SRI randomized one of each pair of schools into treatment and the other of the pair into the control condition. Such blocked randomization (1) balances grade levels between treatment and control, (2) balances policy context between treatment and control, and (3) allows for the local Writing Project sites to roll out implementation evenly between treatment and control cohorts (e.g., no site ends up with all treatment).

Teachers were recruited within schools by the local Writing Project sites prior to randomization. Sites worked with their partner districts to identify eligible teachers and target them for recruitment. To be eligible, teachers had to have taught ELA or ELD in grades 7–11 and be willing to make a 2-year commitment to the study. In addition, teachers agreed to implement Pathway in their focal classes and not to share Pathway materials with comparison teachers if their schools were randomized into the treatment condition. The randomly assigned school sample consisted of 46 schools with 208 secondary ELA and ELD teachers across grades 7–11.

Prior to school randomization, **SRI chose two focal classes for each teacher**.[4] SRI collected teachers' class assignments and chose the focal classes from eligible classes on these

[4] In a few cases, teachers taught only one class that met the inclusion criteria described above.

lists. Eligible focal classes were ELA and ELD classes in grades 7–11, excluding any classes for which the writing assessment would be inappropriate (e.g., self-contained special education or ELD classes for newly arrived immigrant students), as well as elective (e.g., creative writing or journalism) and support (e.g., prep for a high school exit exam) courses. In selecting focal classes, where possible, we avoided AP and IB classes as these classes are prescriptive in terms of the material that needs to be covered, and we chose two identical focal classes so that teachers could coordinate their preparation and instruction. Finally, we sought to achieve comparability across the paired schools (e.g., if a teacher taught eligible classes at more than one grade level, we selected the grade level that best aligned with the focal-class grade level at the paired school).

The evaluation was conducted over 3 years, with local Writing Project sites and their partner districts launching the Pathway program at three different points in time, resulting in three cohorts of districts/schools: fall 2019 (Cohort 1), fall 2020 (Cohort 2), and fall 2021 (Cohort 3). Cohort 1 and 2 schools participated in the program and evaluation for 2 years, while Cohort 3 schools participated for just 1 year. As shown in Exhibit 2, Cohort 1 and Cohort 2 schools had three types of students:

- Group A: Year 1-only students
- Group B: Students who were in a focal class in both Year 1 and Year 2
- Group C: Year 2-only students (i.e., students in a focal class for the first time in Year 2)

Cohort 3 schools had only Year 1 students (Group A).

Exhibit 2. District/School Cohorts and Student Groups

<table>
<tr><th>Group</th><th colspan="3">Year 1</th><th colspan="3">Year 2</th></tr>
<tr><td>A (Year 1 only)</td><td>Baseline</td><td>Grade 7–11 focal class students</td><td>Outcome</td><td colspan="3">Leave school (including move to grades 9 & 12)
Not in focal class</td></tr>
<tr><td>B (Years 1 & 2)</td><td>Baseline</td><td>Grade 7, 9 & 10 focal class students</td><td></td><td>Grade 8, 10 & 11 focal class students</td><td>Outcome</td></tr>
<tr><td>C (Year 2 only)</td><td colspan="3">Not in focal class</td><td>Baseline</td><td>Grade 7–11 focal class students</td><td>Outcome</td></tr>
</table>

Note. For Cohort 1, 2, and 3 schools, Year 1 is 2019–20, 2020–21, and 2021–22, respectively. Cohort 3 schools only have 1 year of the intervention and therefore do not have Group B or C students.

In Year 1, the student sample was defined by enrollment in a focal class. For the writing outcomes analysis, student sample inclusion was based on class rosters collected before randomization or pre-randomization writing assessments administered by SRI.[5] In two sites, some or all classes were scheduled in half-year blocks (i.e., courses were offered in either the fall

[5] If the administration window for the pretest was prior to randomization, we used the pretest submitted prior to randomization to determine which students were enrolled in focal classes. If the administration window for the pretest was after randomization, we used district-provided class rosters collected prior to randomization to determine which students were enrolled in focal classes.

semester or the spring semester). In these sites, the baseline assessment was administered midyear, after randomization was announced.[6] Therefore, the student sample was defined based on spring class rosters, built at the same time as fall semester class rosters and collected prior to randomization, and not based on participation in the writing assessment. In the remaining sites, the writing assessments were administered prior to randomization, so the assessment submission was used to determine focal class enrollment.

For the ELA achievement analysis, Year 1 student sample inclusion was based on focal class rosters collected before randomization, regardless of whether a student took the writing assessment. In those schools in which classes were offered in half-year blocks, only students enrolled in spring semester classes were included in the analysis.

In Year 2, Group B ("continuing") and Group C ("first-timer") students were placed into classes by school counselors who were aware of both teachers' and continuing students' treatment conditions. SRI provided rosters of Group A students who were in focal classes in Year 1 to aid in Year 2 class placement. For the Group B student sample, these counselors had to balance their charge to maintain intact treatment conditions for the sake of the evaluation with competing scheduling demands. Thus, the Year 1 sample consisted of randomly assigned students. The Year 2 Group B sample consisted of Year 1 randomly assigned students but had high attrition by design, and the Year 2 Group C sample consisted of in-moving students.

Independence of the Impact Evaluation

The SRI researchers who conducted this evaluation were external to and independent of UCIWP. SRI worked closely with the intervention team at UC Irvine to conceptualize the study, ensuring that we captured the key program components in our implementation study and designed an impact study likely to capture the impacts of the program (e.g., by selecting independent but aligned outcome measures and advising on school recruitment given the estimated minimally detectable effect size), if such impacts existed.

Once designed, SRI independently conducted all key aspects of the evaluation described in this report, including collection of class rosters to determine analytic samples, random assignment, collection of key outcomes data (other than from administrative records), analyses, and reporting of study findings. We consulted the UCIWP implementation team regarding all key design decisions throughout the study to ensure any changes maintained the intended alignment to the program as designed. The UCIWP team was also given a chance to review this report before it was finalized and suggest changes or edits. However, neither individual decisions nor the final content of this report were subject to the approval of the project director or staff who conceptualized and implemented the intervention.

[6] We included only students enrolled in spring semester classes, which allowed us to assess impacts on students when their teachers had access to a full year of professional development.

Preregistration of the Study Design

Following EIR grant requirements, the SRI research team preregistered this study design. Specifically, the grant's evaluation technical assistance team required all evaluators to preregister design, data, and analysis plans for confirmatory (i.e., primary) impact questions. After the grant's technical assistance team reviewed a detailed design plan to ensure the study was likely to meet grant requirements, SRI then registered the design in the REES (Study Registry ID 3520), a database of impact studies in education (and related fields) funded by IES. We originally registered the study on May 28, 2021, which classified the study as a cluster RCT, following an analysis plan that prioritized the study's ability to meet WWC standards without reservations. Over the next year, the research team recognized that (1) overall attrition would be high enough to classify the RCT as "meets What Works Clearinghouse standards with reservations" and (2) the treatment was unlikely to be delivered with fidelity because of the COVID-19 pandemic. In particular, we believed that this compromised implementation blurred meaningful distinctions between students and teachers being in their first or second year of the program, as teachers were unlikely to receive the intended yearlong dosage in any of the study years. We therefore made changes to the registry on September 2, 2022, to better align the analysis plan to our understanding of the implementation, as delivered. All impact estimates reported in the body of this technical report are included in the preregistered analysis plan.

Changes to Initial Registration in Fall 2022

Our fall 2022 updates to the design plans both relaxed rules on the student analytic samples (originally intended to support the study's ability to meet WWC standards without reservations) and moved the analysis from analyzing separate impact estimates based on the dosage received (i.e., separate estimates for student Groups A, B, and C) by including all study students in a single analysis with statistical controls for student groups. Specifically, we relaxed our rules on joiners and included students in the Year 2 analytic sample even if they were not present in a focal classroom at randomization (i.e., Group C students). Relaxing our approach to joiners allowed us to combine the three distinct groups of students (A, B, and C). Given the impact of pandemic-related disruptions on Pathway implementation in each year of the evaluation, we believe that estimating a single effect across these three groups is the strongest analysis plan. Moreover, we had previously designated any analysis exploring the differences between Groups A, B, and C as exploratory, to avoid overinterpreting potentially spurious distinctions.

We also updated the "type of intervention" question in Section II, Description of Study. As Pathway is a complex intervention with multiple components, this change did not reflect a change in understanding of the program, but merely a change in understanding how the data from this registry might be used. Finally, we updated the logic model, although only to reflect the pandemic-affected context.

Both the first registration and the subsequent edits were made before SRI had access to student outcome data.

Changes in Analysis Since Preregistration

Because of uneven implementation that was exacerbated by the pandemic, SRI opted not to explore differential impact by district or the relationship between fidelity of implementation and effects on student outcomes.

More details on the registered design can be found in the REES document, which is permanent and not editable (https://sreereg.icpsr.umich.edu/sreereg/subEntry/14840/pdf).

Measures

Analytic Writing Assessment

SRI administered a 2-day analytic writing assessment to all students in focal classrooms in the fall of Year 1 (baseline) and again in the spring of Year 1 (outcome), fall of Year 2 (baseline for first-timers only) and spring of Year 2 (outcome). In each year, the specific writing prompts were counterbalanced such that all teachers in a single randomization block were assigned one of two prompts in the fall and the other in the spring.

On Day 1, the prompts provided students with literary nonfiction text to read and several activities designed to scaffold their analysis of the text. On Day 2, students were asked to write an analysis of the text. One set of two prompts was administered in Year 1, and another set of two prompts was administered in Year 2, with the Year 2 prompts having a different design (e.g., to prevent a feeling of repetition for both teachers and students). In Year 1, students received a single text (a short piece of literary nonfiction) and were asked to analyze it for theme and author's craft. In Year 2, students received two texts (a short literary biography and a description of leadership traits). They were asked to write an analytic essay drawing from both texts (e.g., Which elements of leadership did Harriet Tubman exemplify?). Both the analytical reading and the text-based writing are aligned with Common Core State Standards, and the on-demand performance tasks are similar to those included on some state assessments (e.g., Smarter Balanced ELA/literacy).

SRI collected all student writing and prepared the papers for scoring by (1) de-identifying so that student, teacher, school, timepoint (fall/spring), and treatment status was not identifiable and (2) removing unscorable (e.g., blank) papers. The student writing was scored with the Analytic Writing Continuum for Literary Analysis (AWC-LA). More than a decade ago, the National Writing Project developed the Analytic Writing Continuum (AWC), which has been shown to be a valid and reliable measure of student writing (Bang, 2013). The original version of the AWC had been used primarily to score writing rooted in students' personal experiences and therefore did not explicitly measure literary analysis. Members of the National Writing Project's national

leadership team[7] worked with a panel of writing assessment experts to modify the AWC to score literary analysis more accurately, with a focus on the development of English learners' writing. The same panel of writing assessment experts selected and annotated anchor papers to be used in training. The revisions to the AWC and the development of annotated anchor papers were intended to help make explicit for scorers how well-established attributes of effective writing are evident in literary analysis. The resulting AWC-LA retains a structure rooted in the "6+1 Traits" of writing (Culham, 2003) but concentrates on the attributes of literary analysis. Each paper was given a holistic rating as well as ratings on each of four attributes: content, structure, stance, and conventions. Additionally, to assess writing productivity, or whether Pathway helped students generate more writing, we used a word count from the essays.

Most prior studies of Pathway have been conducted by UCIWP and used UCIWP's own rubric, the Assessment of Literary Analysis (ALA) to measure impacts (e.g., Kim et al., 2011; Olson et al., 2012; Olson & Land, 2007). The ALA was designed to be closely aligned to both the Pathway intervention and the prompt itself (e.g., including prompt fulfillment as a facet of the scoring system). By moving away from a program and task-aligned measure and instead using a general measure of writing quality, the AWC-LA allows for a more general interpretation of Pathways impacts (i.e., impacts on students' literary analysis performance, as opposed to performance on the individual performance task; Kane, 2013). As expected from this less aligned and more independent measure (Wolf & Harbatkin, 2023), on SRI's prior independent evaluation of Pathway we found that estimates of Pathway's impacts on student achievement using the AWC-LA resulted in smaller estimated effect sizes than analyses using the ALA (Arshan & Friedrich, 2017).

The AWC-LA rubric uses a 1–6 point scale, with 6 as a high score. Scorers are non-study teachers trained on the AWC-LA rubric and scoring procedure. Scorers first assign scores on each of four individual attributes:

- **Content**: the extent to which the piece presents a compelling claim and supports the thesis using evidence from the text
- **Structure**: the extent to which the writing's structure established an order and arrangement to support the thesis
- **Stance**: the extent to which the writing establishes perspective through tone or style
- **Conventions**: the extent to which the writing demonstrates usage of conventions (e.g., punctuation, spelling) in line with grade-level standards

[7] The National Writing Project is a decentralized network of approximately 170 university-based local Writing Project sites. Local Writing Project sites (including UCIWP) are headed by autonomous university faculty. The national leadership team at the National Writing Project provides support and organizational capacity to facilitate relationships and collaboration among local sites (Lieberman & Wood, 2002). The team also provides writing assessment services for both research and professional development on a fee basis. Pathway was developed at UCIWP, without support from the national leadership team; we therefore consider the AWC-LA to be an independent measure, despite the working relationship between the two autonomous organizations.

Following the scoring of each individual attribute, scorers assign the writing a holistic score. While the holistic score is not formally weighted between the attributes, the holistic rubric describes the score in a way that suggests the content attribute should be most heavily weighted.

The National Writing Project scores by individual attribute, and then scores holistically not to collect the individual attribute scores but to support the most accurate holistic score possible. Scoring by attribute ("primary trait scoring") has been found to produce the more accurate and reliable scores than scoring holistically (Hunter et al., 1996). For example, students' writing may display stronger mastery of content than conventions; this may be particularly true of multilingual students (Bacha, 2001; see also Bang, 2013, for a description of the National Writing Project's scoring framework and process).

We therefore used the holistic score as our primary measure of writing quality. In line with the WWC (2021) Study Review Protocol (Version 1.0) that was available when we preregistered our study, we also looked to the conventions measure as a measure of writing conventions and the word count of each essay as a measure of writing productivity.

Reliability of the prompt scoring was assessed separately for each writing attribute measure in the AWC-LA through the double scoring of a subset of papers. SRI randomly selected 935 papers to be double-scored, and we used the percent-within-one approach (Allen et al., 2013; McCaffrey et al., 2015) to calculate rater agreement for each attribute. Raters agreed within a single score point for 92% of papers on the holistic score, 91% on the content attribute, 90% on the structure attribute, and 89% on the stance and conventions attributes.

We used the Natural Language Toolkit (NLTK) in Python to conduct word count for the essays. The word count excludes infrequent words that are no longer than two letters and short words that are frequently used but often do not contribute much meaning to the content of the text, such as "the," "and," "is," and "in."

State ELA Assessments

As an additional measure of student achievement, we collected state ELA/reading assessment data as available. The state assessment data we used at the middle school level (grades 6 to 8) were the SBAC ELA assessment in California, RISE ELA assessment in Utah, State of Texas Assessments of Academic Readiness (STAAR) ELA, Minnesota Comprehensive Assessments (MCA-III) Reading, Wisconsin Forward Exam ELA, ACT Reading in Arizona, and Oklahoma School Testing Program (OSTP) ELA. At the high school level (grades 9–11), students were assessed on one of the following tests: ACT reading at grade 11, Aspire reading or ELA at grades 9 and 10, SBAC ELA at grade 11 in California, and EOC English in grades 9 and 10 in Texas. For ACT and Aspire assessments, the reading assessment is more aligned with the intervention than the English assessment or the overall ELA score; as a result, we took as the outcome ACT/Aspire reading score where available, and ELA score where the reading score was not available.

In accordance with best practices from the U.S. Department of Education's National Center for Education Evaluation and Regional Assistance (May et al., 2009), we converted student baseline and outcome ELA test scores into standardized *z*-scores for the same test separately for each grade, year, and site, and conducted the impact analysis combining all sites. We used districtwide means and standard deviations for standardizing scores.

Sample Sizes and Attrition

A total of 46 schools were randomized, with 23 to the treatment condition and 23 to the control condition. No school dropped out from the writing analytic sample.

Because different criteria were used to establish the student samples for the writing outcomes analysis and the ELA assessment analysis, we discuss the two analyses separately below.

Student Writing Assessment Sample

For an impact analysis making use of the intact randomization design, we would only include the Year 1 student sample. However, the relatively high attrition rates prevented an analysis of only Year 1 students from providing the highest validity from the RCT design (see Exhibit A-1 in the Appendix). As a result, we did not pursue a separate Year 1 impact analysis for students at the time of randomization.

Because Pathway is designed as a 2-year intervention, for the confirmatory analysis, we ideally would have included the impact of the intervention in Year 2 (for Group C, under teachers who were receiving the second year of intervention) or examine student progress over 2 years (for Group B). However, if we only looked at Year 2 or 2-year impact, we would not be able to include the nine Cohort 3 schools that participated in the intervention for just 1 year. Further, given pandemic-related disruptions to training, schooling, and implementation of the program, we did not believe that this analysis would accurately reflect a distinction between the first and second year of exposure to the intervention for either teachers or students under typical conditions. We therefore decided to include all three groups of students in the confirmatory impact analysis in order to maximize our use of available data, achieve maximum statistical power, and avoid overinterpreting dosage distinctions. We registered this design decision with the REES before we collected outcome data.

The impact analysis only included students who had both baseline and outcome writing assessment scores. There were 2,641 Group A (Year 1 only), 580 Group B (Years 1 and 2), and 1,644 Group C (Year 2 only) students with writing assessment scores. Because of limited capacity for scoring, we took all 580 Group B students and randomly sampled 83% of Group A and C students for scoring, resulting in a final sample of 2,192 Group A, 580 Group B, and 1,364 Group C students from all 46 study schools for the impact analysis. Exhibit 3 displays the number of students and schools by treatment status for the analysis, and Exhibit 4 shows the number of students in each group by treatment condition.

Exhibit 3. Number of Schools and Students in Analytic Sample, Student Writing Outcomes

Outcome Measure	Comparison Group		Treatment Group	
	Clusters	Students	Clusters	Students
Holistic	23	1,904	23	2,232
Conventions	23	1,904	23	2,232
Word count	23	1,904	23	2,232

Exhibit 4. Number of Students in Analytic Sample by Group, Student Writing Outcomes

Group	Comparison	Treatment	Total
Group A (Year 1 only)	1,059	1,133	2,192
Group B (Years 1 and 2)	217	363	580
Group C (Year 2 only)	628	736	1,364
Total	1,904	2,232	4,136

State ELA Assessment Sample

SRI's analysis of state test score data included all three groups of students in all randomly assigned treatment and comparison schools that had both baseline and outcomes assessment scores.[8] Because of the pandemic, none of the sites assessed students in 2019–20, so we relied on 2018–19 ELA assessment data as baseline for cohorts of students who started the intervention in fall 2020–21. For students who started the intervention in fall 2021–22, we used 2020–21 ELA assessment data as baseline. Further, entire districts and grade levels were excluded from the analysis of state test scores because of the lack of available data.[9]

The final analytic sample for ELA achievement included a total of 35 schools, 17 in the treatment and 18 in the comparison condition. There were 2,358 Group A (Year 1 only) students, 941 Group B (Year 1 and 2) students, and 1,736 Group C (Year 2 only) students. Exhibit 5 displays the number of students in each group by treatment condition. All students included in the analysis had both baseline and outcome state ELA/reading assessment scores.

[8] Like the student writing assessment outcomes analysis, the randomized Year 1 student sample had high attrition rates and, as a result, could not provide the highest validity from the RCT design; we therefore did not conduct a separate Year 1 student analysis.

[9] Group A students in Long Beach Unified School District were excluded from the analysis because there was no outcome measure from 2019–20. Los Angeles Unified School District also did not assess students in 2020–21 (baseline for Cohort 3 sites) or 2019-20, so the whole district was excluded from the analysis because there was no baseline measure. Phoenix Union High School District did not assess students in grade 9, so students in grades 9 and 10 were excluded from the analysis because of a lack of either outcome or baseline data.

Exhibit 5. Number of Students in Each of the Groups in Analytic Sample, ELA Assessment

Group	Comparison	Treatment	Total
Group A (Year 1 only)	1,011	1,347	2,358
Group B (Years 1 and 2)	475	466	941
Group C (Year 2 only)	796	940	1,736
Total students	2,282	2,753	5,035
Total schools	18	17	35

Data Analysis and Findings

Baseline Equivalence: Student Writing Outcomes Sample

We estimated baseline equivalence by comparing means and standard deviations of baseline student writing measures between treatment and comparison groups. We calculated the difference using the same structural model used to predict the outcome scores to estimate the treatment effect with student and school levels, but only controlling for blocking effects and using the same sample weights as used in the impact analysis (described below). Exhibit 6 presents sample sizes, means, and standard deviations of baseline outcome measures for comparison and treatment groups separately; the adjusted treatment-comparison differences from the model; and standardized difference (Hedges' *g*) based on the pooled standard deviation of treatment and comparison groups. All standardized differences in outcomes are smaller than 0.25, establishing baseline equivalence, given that we adjusted statistically for baseline measures in the impact analyses.

Exhibit 6. Baseline Equivalence, Student Writing Outcomes Sample

Outcome Measure	Comparison Group			Treatment Group				
	Sample Size	Mean	SD	Sample Size	Mean	SD	Treatment – Comparison Difference	Standardized Difference
Holistic	1,904	2.87	1.25	2,232	2.75	1.25	-0.12	-0.10
Conventions	1,904	3.02	1.30	2,232	2.91	1.31	-0.12	-0.09
Word count	1,904	133.64	92.40	2,232	127.37	85.19	-6.27	-0.07

We further compared student characteristics between the treatment and comparison students (Exhibit 7 and Exhibit 8). Few students were missing student demographic data other than eligibility for free or reduced-price lunch (FRPL), which two districts were not able to provide. As result, we included in the analysis all students with complete demographic data, except FRPL. Students in treatment and comparison groups were very similar across all demographic characteristics as well as student grade level, with the overall differences all smaller than 5 percentage points.

Exhibit 7. Student Demographics, Student Writing Outcomes Sample

Condition	Percent/*n*	Female	Asian	Black	Latino	White	FRPL	EL	IEP
Comparison	Percent	54%	6%	11%	43%	35%	51%	5%	5%
	Total *n*	1,904	1,904	1,904	1,904	1,904	1,587	1,904	1,904
Treatment	Percent	51%	7%	9%	45%	33%	49%	9%	7%
	Total *n*	2,232	2,232	2,232	2,232	2,232	1,794	2,232	2,232

Note. FRPL = free or reduced-price lunch; EL = English learner; IEP = Individualized Education Program.

Exhibit 8. Student Grade Level When Entering the Intervention, Student Writing Outcomes Sample

		Entering Grade					
Condition	Percent/*n*	7	8	9	10	11	Total
Comparison	Percent	30%	22%	20%	20%	8%	100%
	n	578	414	380	383	149	1,904
Treatment	Percent	35%	18%	24%	17%	6%	100%
	n	789	401	537	381	124	2,232

Program Effects

We used a two-level hierarchical linear model (HLM) with student and school levels to examine the impact of Pathway on student writing outcomes, as below:

$$Y_{ij} = \beta_0 + \beta_1 (Pathway_j) + \beta_k (kth - student\ covariate_{ij}) + \beta_l (lth - school\ covariate_j) + \beta_m (mth - randomization\ block_j) + e_{ij} + r_j$$

Impacts were estimated separately for each writing outcome score. Random effects e_{ij} and r_j allow for error at the student and school levels. The model was adjusted for student (e.g., baseline outcome score, grade level, and demographic information) and school (school aggregate demographic information) covariates, as well as randomization block effects. (See Exhibit A-2 in the Appendix for a list of student and school covariates included in the model.)

Because we sampled a larger proportion of eligible Group B students (100%) than the other two groups (83%), we weighted the analytic sample to make it representative of all students who had both baseline and outcome writing assessments. Because of a lack of reliable FRPL data in two districts, 18% of students (17% in comparison and 20% in treatment) in the analytic sample were missing FRPL information. We applied the dummy variable method to handle missing values on FRPL, where we set missing values for FRPL to the sample mean and created a dummy variable indicator with a value of 1 for students missing FRPL and a value of 0 for students without missingness.

Exhibit 9 has descriptive information and impact estimates for the student writing outcomes. The Pathway program increased the performance of students in the treatment group 0.12 standard deviations more than that of students in the comparison group ($p < .01$) on the holistic

score. Results were also positive and statistically significant for convention score and word count, with effect sizes of 0.10 and 0.19 respectively. (The AWC-LA scoring system also generates ratings for three additional writing attributes: content, structure, and stance; the descriptives and estimated program impact on these attributes are presented in Exhibits A-3 and A-4 in the Appendix.)

Exhibit 9. Descriptives of Outcome Scores and Estimated Program Impact, Student Writing Assessment

Outcome Measure	Comparison Group			Treatment Group			Treatment–Comp. Difference	*SE*	Standardized Difference	*p*-Value
	Sample Size	Mean	*SD*	Sample Size	Model-Adjusted Mean	*SD*				
Holistic	1,904	3.34	1.34	2,232	3.50	1.39	0.16	0.06	0.13	.009
Conventions	1,904	3.44	1.32	2,232	3.58	1.37	0.14	0.05	0.11	.005
Word count	1,904	166.24	99.52	2,232	186.32	112.52	20.08	5.59	0.23	.000

Note. *SD* = standard deviation; *SE* = standard error.

To further test if the impact of Pathway on student writing quality differs by student baseline characteristics (gender, FRPL, eligibility for English learner services) or student cohort groups (Groups A, B, and C), we added interaction terms between each of the characteristics and the treatment indicator to the holistic score impact model. None of the interaction terms were statistically significant, indicating a lack of differential program impact on student subgroups. This uniformity in impact across diverse student demographics underscores the intervention's potential to provide equitable learning opportunities and outcomes for all participants.

Baseline Equivalence: State ELA Assessments

We estimated baseline equivalence by comparing means and standard deviations of baseline ELA scores between treatment and comparison groups. We calculated the difference using the same structural model used to predict the outcome scores to estimate the treatment effect with student and school levels, but only controlling for blocking effects. Exhibit 10 presents sample sizes, means, and standard deviations of baseline ELA scores for comparison and treatment groups separately, the adjusted treatment-comparison difference from the model, and standardized difference (Hedges' *g*) based on the pooled standard deviation of treatment and comparison groups. The standardized differences in baseline ELA scores is 0.11. much smaller than 0.25, establishing baseline equivalence, given that we adjust statistically for the baseline measure in the impact analysis.

Exhibit 10. Baseline Equivalence for State ELA Assessment Sample

Comparison Group			Treatment Group				
Sample Size	Mean	*SD*	Sample Size	Mean	*SD*	Treatment–Comparison Difference	Standardized Difference
2,282	-0.03	0.88	2,753	0.07	0.90	0.10	0.11

Note. *SD* = standard deviation.

As shown in Exhibits 11 and 12, students in treatment and comparison groups were also very similar on all demographic characteristics and grade-level distribution, with overall differences all smaller than 5 percentage points except for the percentage of Latino students, which was 35% in the treatment condition and 25% in the comparison condition.

Exhibit 11. Student Demographics, State ELA Assessment Sample

Condition	Percent/ *n*	Female	Asian	Black	Latino	White	FRPL	EL	IEP
Comparison	Percent	48%	4%	14%	25%	6%	43%	7%	9%
	Total *n*	2,282	2,282	2,282	2,282	2,282	1,999	2,282	2,282
Treatment	Percent	47%	4%	10%	35%	4%	41%	8%	9%
	Total *n*	2,753	2,753	2,753	2,753	2,753	2,376	2,753	2,753

Note. FRPL = free or reduced-price lunch; EL = English learner; IEP = Individualized Education Program.

Exhibit 12. Student Grade Level When Entering the Intervention, State ELA Assessment Sample

Condition	Percent/ *n*	Entering Grade					Total
		7	8	9	10	11	
Comparison	Percent	46%	31%	13%	4%	7%	100%
	n	1,053	697	291	88	153	2,282
Treatment	Percent	49%	26%	11%	7%	6%	100%
	n	1,343	726	314	201	169	2,753

Program Effects

As described earlier, we used a two-level HLM with student and school levels to examine the impact of Pathway on student ELA/reading achievement. Because of a lack of reliable FRPL data in two districts, 13% of students (12% in comparison and 14% in treatment) in the analytic sample were missing FRPL information. We applied the same dummy variable method to deal with missing values on FRPL as in the writing outcomes analysis.

Exhibit 13 has descriptive information and impact estimate for the state ELA assessment. The estimated impact is not statistically significant (p = .61).

Exhibit 13. Descriptives of Outcome Scores and Estimated Program Impact, State ELA Assessment Sample

Comparison Group			Treatment Group						
Sample Size	Mean	*SD*	Sample Size	Model-Adjusted Mean	*SD*	Treatment–Comparison Difference	*SE*	Standardized Difference	*p*-Value
2,282	-0.14	0.91	2,753	-0.13	0.90	0.01	0.02	0.01	.610

Note. *SD* = standard deviation; *SE* = standard error.

Fidelity of Implementation Study

SRI researchers worked collaboratively with Pathway program developers at UCIWP to define the key program components, specify indicators for each component, and design an approach to measure each indicator. In this section, we describe our approach first and then present implementation study findings.

Fidelity Measurement

To measure fidelity of implementation (FOI), SRI researchers worked with UCIWP program developers to specify indicators for each key program component included in the Pathway logic model (Exhibit 14). For each indicator, we established the appropriate unit of implementation, thresholds for adequate implementation, and aligned data sources.[10] At the component level, we specified a definition of adequate implementation based on the percentage of local Writing Project sites reaching the specified indicators; for each component, we set the adequacy threshold at 100% of local Writing Project sites reaching each individual indicator threshold.

Exhibit 14. Scoring System for Fidelity of Implementation of Each Key Component in the Pathway Logic Model

Component 1: Professional Development Content			
Key Indicators	**Unit of Implementation**	**Threshold**	**Data Source**
Indicator 1.1. Site presents teacher participants with cognitive strategies reading tutorial	Writing Project site	90% of teachers attend cognitive strategies reading tutorial PD	Attendance trackers and artifact analysis
Indicator 1.2. Site presents teacher participants with cognitive strategies writing tutorial	Writing Project site	90% of teachers attend cognitive strategies writing tutorial PD	Attendance trackers and artifact analysis
Indicator 1.3. Site presents teacher participants with the book club binder (Year 1 only)	Writing Project site	90% of teachers attend book club tutorial PD	Attendance trackers and artifact analysis
Indicator 1.4. Site presents teacher participants with the revision tutorial	Writing Project site	90% of teachers attend revision tutorial PD	Attendance trackers and artifact analysis
Indicator 1.5. Site asks teacher participants to take students through process of revising their pretest	Writing Project site	90% of teachers attend PD where site asks teacher participants to take students through process of revising their pretest	Attendance trackers and artifact analysis

[10] For Key Components 1 and 2, the unit of implementation, threshold, and data source are the same because the indicators are based on teacher attendance at professional development events that incorporated both content and a scaffolded approach.

Component 2: Scaffolded Approach			
Key Indicators	**Unit of Implementation**	**Threshold**	**Data Source**
Indicator 2.1. Site presents model lessons showcasing Pathway strategies from cognitive strategies reading tutorial	Writing Project site	90% of teachers attend cognitive strategies reading tutorial PD	Attendance trackers and artifact analysis
Indicator 2.2. Site presents model lessons showcasing Pathway strategies from cognitive strategies writing tutorial	Writing Project site	90% of teachers attend cognitive strategies writing tutorial PD	Attendance trackers and artifact analysis
Indicator 2.3. Site presents model lessons showcasing Pathway strategies from book club binder (Year 1 only)	Writing Project site	90% of teachers attend book club tutorial PD	Attendance trackers and artifact analysis
Indicator 2.4. Site presents model lessons showcasing Pathway strategies from revision tutorial	Writing Project site	90% of teachers attend revision tutorial PD	Attendance trackers and artifact analysis
Indicator 2.5. By the time teachers attend revision tutorial PD, site distributes key project materials to teacher participants (i.e., cognitive strategies bookmarks, wall poster, blue booklets, book club binder, cognitive strategies tutorial PowerPoint for teachers, teaching theme tutorial PowerPoint for teachers, teacher PD videos, student mini-tutorial videos, and revision task tutorial PowerPoint for teachers)	Writing Project site	Once per site	Artifact analysis
Component 3: Formative Feedback			
Key Indicators	**Unit of Implementation**	**Threshold**	**Data Source**
Indicator 3.1. Site reads and provides feedback on all treatment students' pretest writing (treatment students = those students in sample at baseline, still present in the spring semester)	Writing Project site	90% of teachers report receiving feedback on at least most (more than 50%) students' pretest papers	Teacher survey
Indicator 3.2. Site prepares and distributes an individualized teacher results letter to all participating treatment teachers (Year 2 only)	Writing Project site	90% of teachers report receiving a teacher results letter	Teacher survey

Component 4: Classroom Implementation			
Key Indicators	**Unit of Implementation**	**Threshold**	**Data Source**
Indicator 4.1. Treatment teachers teach Pathway introductory reading tutorial in two focal classes	Writing Project site	90% of teachers teach all or part of the Pathway reading introductory tutorial	Teacher survey
Indicator 4.2. Treatment teachers teach Pathway introductory writing tutorial in two focal classes	Writing Project site	90% of teachers teach all or part of the Pathway writing introductory tutorial	Teacher survey
Indicator 4.3. Treatment teachers teach Pathway revision tutorial in two focal classes	Writing Project site	90% of teachers teach all or part of the Pathway revision tutorial	Teacher survey
Indicator 4.4. Treatment teachers engage students in multiple-draft essay based on reader comments	Writing Project site	90% of teachers engage students in multiple-draft essay	Teacher survey

Note. PD = professional development.

Fidelity Findings

To assess FOI across the participating local Writing Project sites, SRI researchers relied on program attendance data, program artifacts (e.g., agendas and presentations), and teacher survey data. SRI collected attendance data and program artifacts twice a year, in January and June, over the 2 years of program implementation. SRI administered a teacher survey in the spring of each year, near the conclusion of the school year. SRI researchers then used this information to assess implementation at the Writing Project site level. For example, for the FOI indicators for Component 1 (professional development content), SRI researchers reviewed artifacts from each professional development event to determine if and when the content was offered and then examined teacher attendance data to assess the percentage of teachers in attendance at the corresponding event. In this case, we found that all sites offered the expected content, but sites did not reach implementation thresholds because they did not achieve 90% teacher attendance (Exhibit 15). We used the same method for Component 2 (scaffolded approach) and found the same result. For Component 3 (formative feedback) and Component 4 (classroom implementation), we relied on teacher surveys. Because of pandemic-related disruptions, the program did not reach the thresholds established for implementation fidelity for any of the four components.

Exhibit 15. Findings on Fidelity of Implementation by Component Over 2 Years

Key Components, Number of Indicators, Units, and Threshold				Year 1 Results			Year 2 Results		
Key Component	Total # of Measurable Indicators	Unit of Implementation	Sample-Level Threshold for Fidelity of Implementation	# of Units in Which Component Was Implemented	# of Units in Which Fidelity of Component Was Measured	% Achieved Fidelity Score and Whether Program Met Sample-Level Threshold	# of Units in Which Component Was Implemented	# of Units in Which Fidelity of Component Was Measured	% Achieved Fidelity Score and Whether Program Met Sample-Level Threshold
1. Professional development content	5 in Year 1 4 in Year 2	WP site	100% of WP sites reaching threshold on all indicators	1 program 8 WP sites	1 program 8 WP sites	37.5% of WP sites ***Program fidelity = No***	1 program 6 WP sites	1 program 6 WP sites	33.3% of WP sites ***Program fidelity = No***
2. Scaffolded approach	5 in Year 1 4 in Year 2	WP site	100% of WP sites reaching threshold on all indicators	1 program 8 WP sites	1 program 8 WP sites	37.5% of WP sites ***Program fidelity = No***	1 program 6 WP sites	1 program 6 WP sites	33.3% of WP sites ***Program fidelity = No***
3. Formative feedback	1 in Year 1 2 in Year 2	WP site	100% of WP sites reaching threshold on all indicators	1 program 8 WP sites 105 teachers	1 program 8 WP sites 100 teachers	12.5% of WP sites ***Program fidelity = No***	1 program 6 WP sites 89 teachers	1 program 6 WP sites 72 teachers	33.3% of WP sites ***Program fidelity = No***
4. Classroom implementation	4 in Year 1 4 in Year 2	WP site	100% of WP sites reaching threshold on all indicators	1 program 8 WP sites 105 teachers	1 program 8 WP sites 100 teachers	0% of WP sites ***Program fidelity = No***	1 program 6 WP sites 89 teachers	1 program 6 WP sites 72 teachers	16.6% of WP sites ***Program fidelity = No***

Note. WP = Writing Project.

In addition to assessing implementation fidelity, the study examined the contrast between teachers' professional development experience in both treatment and comparison schools via a teacher survey. Administered in the late spring of each year, the survey included questions about the duration and content of all ELA-focused professional development received during the evaluation. Exhibits 16 through 20 present the results across all schools during their first year of program implementation (2019–20 for Cohort 1, 2020–21 for Cohort 2, and 2021–22 for Cohort 3). Exhibits 21 through 25 show the Year 2 results for the subset of schools in which teachers participated in 2 years of Pathway professional development (2020–21 for Cohort 1 and 2021–22 for Cohort 2). In both Years 1 and 2, Pathway teachers received an average of 18 hours more professional development than comparison teachers (28 hours compared with 10 hours; see Exhibits 16 and 21). Moreover, in both years, Pathway teachers generally reported that the ELA-focused professional development they participated in was aligned with their classroom needs, such as relating to the standards they teach, aligning to their state assessments, and applying to the needs of their students (see Exhibits 17 and 22). Pathway teachers also reported that the ELA-focused professional development they participated in had greater emphasis than comparison teachers' professional development on cognitive strategies in reading (e.g., making predictions about what will happen in a text, making connections with the author or a character in the text; see Exhibits 18 and 23) and writing (e.g., balancing the use of summary, supporting detail, and commentary; composing a thesis statement or developing a claim; sentence craft and variety; see Exhibits 19 and 24). Finally, Pathway teachers were more likely than comparison teachers to report that the ELA-focused professional development they participated in included the scaffolding that is central to the Pathway model, such as providing teachers with prepared lessons to support classroom implementation and analyzing student work to inform instruction (see Exhibits 20 and 25).

Exhibit 16. Amount of Professional Development, Year 1

Approximately how many hours of ELA-focused professional development have you had this school year?				
	Treatment	Comparison		Teacher *N*
Hours of writing professional development received	28.2	10.0	***	193

*** $p < .001$.

Exhibit 17. Professional Development Alignment to Classroom Needs, Year 1

To what extent did the ELA-focused professional development align to your classroom's needs:				
	Treatment	Comparison		Teacher *N*
Related to or aligned with my state's standards	2.84	2.69	*	171
Related to or aligned with my school or district's ELA curriculum	2.63	2.72		169
Related to or aligned with my school or district's pacing guide	2.26	2.45		159
Aligned with the state assessment my students will take this year	2.62	2.28	**	159
Aligned with the benchmark assessments used in my district	2.54	2.36		159
Aligned with my personal goals for my students' learning	2.77	2.58	*	168
Focused on the needs of English learners	2.63	2.31	**	167
Focused on the needs of students who read below grade level	2.60	2.20	***	170
Focused on the needs of students who write below grade level	2.60	2.23	***	168
Was applicable to all students I teach	2.73	2.47	**	169

Note. Scale: 1 = Not at all, 2 = Minor extent, 3 = Major extent.
* $p < .05$; ** $p < .01$; *** $p < .001$.

Exhibit 18. Professional Development Emphasis on Cognitive Strategies in Reading, Year 1

In the ELA-focused professional development, to what extent were the following reading strategies emphasized:				
	Treatment	Comparison		Teacher *N*
Establishing a goal before reading a particular text	3.09	2.65	**	171
Thinking about or discussing prior knowledge before reading	3.42	2.85	***	170
Making predictions about what will happen in a text	3.45	2.61	***	168
Making connections with the author or a character in a text	3.46	2.84	***	171
Summarizing main ideas or key points	3.58	3.01	***	171
Analyzing the meaning of a text	3.63	2.89	***	171
Identifying evidence to support interpretations of a text	3.62	3.23	**	170
Annotating a text to ask questions, make predictions, and form interpretations	3.45	2.87	***	171
Visualizing characters or scenes in a text to create a mental picture of what is happening	3.16	2.44	***	170
Monitoring one's understanding while reading a text	3.39	2.58	***	170
Asking questions to clarify meaning when understanding breaks down	3.41	2.73	***	171
Analyzing the author's style	3.27	2.60	***	170
Revising one's interpretation of the text throughout the reading process	3.08	2.37	***	170

Note. Scale: 1 = No/ Almost no emphasis, 2 = Minor emphasis, 3 = Some emphasis, 4 = Significant emphasis, 5 = Heavy emphasis.
** $p < .01$; *** $p < .001$.

Exhibit 19. Professional Development Emphasis on Writing Strategies, Year 1

In the ELA-focused professional development, to what extent were the following writing strategies and skills emphasized:				
	Treatment	Comparison		Teacher *N*
Deconstructing a prompt	3.46	2.54	***	170
Prewriting or planning	3.35	2.91	**	170
Organizing a piece of writing	3.47	2.85	***	170
Composing a thesis statement or developing a claim	3.70	2.95	***	170
Using relevant facts, details, quotes, and examples in writing	3.54	2.93	***	170
Sentence craft and variety	2.92	2.26	***	170
Using academic language	3.04	2.55	***	170
Specific writing techniques	3.21	2.57	***	170
Balancing the use of summary, supporting detail, and commentary in an essay	3.38	2.35	***	170

Note. Scale: 1 = No/ Almost no emphasis, 2 = Minor emphasis, 3 = Some emphasis, 4 = Significant emphasis, 5 = Heavy emphasis.
** $p < .01$; *** $p < .001$.

Exhibit 20. Professional Development Characteristics, Year 1

Did the ELA-focused professional development include the following:				
	Treatment	Comparison		Teacher *N*
Modeling instruction with teachers participating as learners	97	69	***	171
Providing teachers with digital materials to support implementation	100	75	***	171
Providing teachers with prepared lessons to implement	90	41	***	171
Analyzing student work to inform instruction	92	52	***	171
Time for discussion of classroom implementation	88	68	***	171

Note. Stems were offered as "check all that apply" and analyzed as binary outcomes using multilevel logistic regression. Outcomes were transformed into percentage points for ease of interpretation and represent the predicted experience for the average teacher in the sample.
*** $p < .001$.

Exhibit 21. Amount of Professional Development, Year 2

Approximately how many hours of ELA-focused professional development have you had this school year?				
	Treatment	**Comparison**		**Teacher *N***
Hours of writing professional development received	28.4	10.3	***	116

*** $p < .001$.

Exhibit 22. Professional Development Alignment to Classroom Needs, Year 2

To what extent did the ELA-focused professional development align to your classroom's needs:				
	Treatment	**Comparison**		**Teacher *N***
Related to or aligned with my state's standards	2.93	2.68	**	99
Related to or aligned with my school or district's ELA curriculum	2.86	2.75		99
Related to or aligned with my school or district's pacing guide	2.56	2.26	**	91
Aligned with the state assessment my students will take this year	2.61	2.18	**	98
Aligned with the benchmark assessments used in my district	2.52	2.20	*	89
Aligned with my personal goals for my students' learning	2.87	2.57	**	99
Focused on the needs of English learners	2.73	2.20	***	100
Focused on the needs of students who read below grade level	2.67	2.22	***	100
Focused on the needs of students who write below grade level	2.73	2.22	***	100
Was applicable to all students I teach	2.89	2.53	**	100

Note. Scale: 1 = Not at all, 2 = Minor extent, 3 = Major extent.
* $p < .05$; ** $p < .01$; *** $p < .001$.

Exhibit 23. Professional Development Emphasis on Cognitive Strategies in Reading, Year 2

In the ELA-focused professional development, to what extent were the following reading strategies emphasized:				
	Treatment	**Comparison**		**Teacher *N***
Establishing a goal before reading a particular text	3.23	2.48	***	102
Thinking about or discussing prior knowledge before reading	3.45	2.85	***	102
Making predictions about what will happen in a text	3.43	2.37	***	102
Making connections with the author or a character in a text	3.55	2.78	***	102
Summarizing main ideas or key points	3.48	2.98	**	102
Analyzing the meaning of a text	3.54	2.76	***	102
Identifying evidence to support interpretations of a text	3.59	3.02	***	102
Annotating a text to ask questions, make predictions, and form interpretations	3.29	2.70	***	102
Visualizing characters or scenes in a text to create a mental picture of what is happening	3.38	2.15	***	102
Monitoring one's understanding while reading a text	3.45	2.57	***	102
Asking questions to clarify meaning when understanding breaks down	3.50	2.54	***	102
Analyzing the author's style	3.20	2.20	***	102
Revising one's interpretation of the text throughout the reading process	3.21	2.24	***	102

Note. Scale: 1 = No/Almost no emphasis, 2 = Minor emphasis, 3 = Some emphasis, 4 = Significant emphasis, 5 = Heavy emphasis.
** $p < .01$; *** $p < .001$.

Exhibit 24. Professional Development Emphasis on Writing Strategies, Year 2

In the ELA-focused professional development, to what extent were the following writing strategies and skills emphasized:				
	Treatment	Comparison		Teacher *N*
Deconstructing a prompt	3.69	2.31	***	100
Prewriting or planning	3.48	2.63	***	102
Organizing a piece of writing	3.46	2.70	***	102
Composing a thesis statement or developing a claim	3.75	2.64	***	101
Using relevant facts, details, quotes, and examples in writing	3.48	2.83	***	102
Sentence craft and variety	2.96	1.98	***	102
Using academic language	3.09	2.28	***	101
Specific writing techniques	3.39	2.30	***	102
Balancing the use of summary, supporting detail, and commentary in an essay	3.55	2.11	***	102

Note. Scale: 1 = No/Almost no emphasis, 2 = Minor emphasis, 3 = Some emphasis, 4 = Significant emphasis, 5 = Heavy emphasis.
*** $p < .001$.

Exhibit 25. Professional Development Characteristics, Year 2

Did the ELA-focused professional development include the following:				
	Treatment	Comparison		Teacher *N*
Modeling instruction with teachers participating as learners	96	71	**	100
Providing teachers with digital materials to support implementation	100	78	***	100
Providing teachers with prepared lessons to implement	93	33	***	100
Analyzing student work to inform instruction	91	29	***	98
Time for discussion of classroom implementation	93	62	***	99

Note. Stems were offered as "check all that apply" and analyzed as binary outcomes using multilevel logistic regression. Outcomes were transformed into percentage points for ease of interpretation and represent the predicted experience for the average teacher in the sample.
** $p < .01$; *** $p < .001$.

Scale-Up Evaluation Study

Strategy to Scale

In the long term, UCIWP envisions leveraging the national network of local Writing Project sites to build durable capacity to implement Pathway across the country. Through this EIR grant, UCIWP sought to (1) expand local Writing Project sites' capacity to implement Pathway and (2) demonstrate that Pathway could be effective in meeting the needs of diverse learners in different state and local contexts. To this end, UCIWP developed scaling strategies to address multiple known barriers to spread and scale: training across geographic distances, meeting the needs of diverse student populations (i.e., students coming from different home language backgrounds), building site capacity, and reducing cost. For each barrier, UCIWP tested strategies to address the challenge.

- To train site leaders across geographic distances, UCIWP invested in national convenings, video recordings of professional development, and technology tools for classroom use.
- To meet the needs of diverse student populations, UCIWP developed mini lessons geared toward different English learner populations.
- To build site capacity, UCIWP invested in training and supporting local site directors and "teacher consultants" who were able to support teachers in their service areas to implement the Pathway model in their classrooms. UCIWP did this through national convenings and by providing support for "teacher consultants" to implement Pathway in their own classrooms. UCIWP also provided "thinking partners" to participating sites to help them strategize about the work with area schools and districts.
- To reduce cost, UCIWP codified instructional materials (e.g., video recordings of exemplar teachers, livestream) and training modules and planned to make materials available at no cost via open-source platforms.[11]

Measurement of Implementation of Scaling Strategies

To assess whether the identified scaling strategies were implemented as intended, the SRI team worked with UCIWP leaders to establish specific thresholds for the implementation of each scaling strategy and to identify appropriate data sources, including convening/meeting artifacts, attendance records, and interviews to determine if a strategy was implemented (Exhibit 26).

[11] Efforts to reduce cost are not discussed here because they were outside the scope of the evaluation and materials were made available after the conclusion of the evaluation.

Exhibit 26. Measurement of Scaling Strategies

Scale-Up Goal	Challenge to Meeting Goal	Strategy to Address Challenge	Definition of Full Implementation of Strategy (in measurable units)	Threshold for Level of Implementation Defined as "Successful"	Data Collection and Reporting Plan for Measuring Implementation of Strategy
Expand to new regions	Geographic distance	National convenings	2 national convenings led by Pathway leadership team by the end of each year (SY 19/20, SY 20/21, SY 21/22)	2 national convenings led by Pathway leadership team by the end of each year (SY 19/20, SY 20/21, SY 21/22)	Review artifacts (i.e., meeting agendas), observe
		Video recordings of UCIWP professional development	All core PD events video-recorded and provided to other WP sites (SY 19/20, SY 20/21)	All core PD events video-recorded and provided to other WP sites (SY 19/20, SY 20/21)	Review artifacts (i.e., videos posted to project website)
		Technology tools for classroom use	Instructional videos developed on 5 topics; sentence fluency tool; revision highlighter	Instructional videos developed on 5 topics; sentence fluency tool; revision highlighter	Review artifacts (i.e., technology tools)
Meet the needs of diverse student populations	Linguistic diversity	Language Support Guidebook	Resource includes salient language features teachers can use to support students' academic writing and is publicly available	Resource includes salient language features teachers can use to support students' academic writing and is publicly available	Review artifacts (i.e., Language Support Guidebook)
		New research	2 journal publications and conference presentations	2 journal publications and conference presentations	Artifacts (i.e., journal articles and presentation slides)
Build WP site capacity	WP site leaders lack experience with and of ownership of the Pathway model	Develop "teacher-consultant" (TC) classroom teachers to serve as Site PD Leads	Each expansion site (7) brings 2 TCs to national convenings (2 trainings/year)	Each expansion site (7) brings at least 1 TC to each national convening, beginning a year in advance of launching the PD with their partner district	Attendance at national convenings
			TCs use Pathway materials in their own classroom during training year	In the year prior to program launch or during Year 1 of implementation, at least one TC from each expansion site	Interviews

Scale-Up Goal	Challenge to Meeting Goal	Strategy to Address Challenge	Definition of Full Implementation of Strategy (in measurable units)	Threshold for Level of Implementation Defined as "Successful"	Data Collection and Reporting Plan for Measuring Implementation of Strategy
				administers pretest and posttests and implements the cognitive strategies and revision tutorials	
		Develop site directors' (SDs') understanding of Pathway model	Each expansion SD attends national convenings (2 trainings/year)	Each expansion SD attends at least one national convening per year	Attendance at national convenings
			SD actively engages with TC to develop programmatic understanding	In the year prior to program launch or during Year 1 of implementation, SD observes TCs' implementation of cognitive strategies and revision tutorials and/or debriefs video of classroom implementation with TCs	Interviews
		Deploy "thinking partners" (TPs) to each expansion site	NWP develops TPs via monthly meetings	85% of the 7 UCI TPs attend at least 7 meetings per year with NWP	Meeting attendance
			TPs support each expansion site implementation team	100% of TPs conduct annual site visit to each expansion site	Meeting minutes

Note. SY = school year; UCIWP = UC Irvine Writing Project; WP = Writing Project; PD = professional development; NWP = National Writing Project.

Findings on Implementation of Scaling Strategies

The UCIWP employed six strategies to support the spread and scale of Pathway to seven new local Writing Project sites in seven states. The strategies were mostly implemented as planned: 6.5 of 8 strategies were fully implemented (Exhibit 27). Where implementation fell short of expected thresholds, the reasons included challenges associated with the COVID-19 pandemic.

Exhibit 27. Overall Findings on Implementation of Scale-Up Strategies at End of Project

Scale-Up Strategy	Threshold for Successful Implementation	Findings on Actual Level of Implementation	Implementation of Strategy Met or Exceeded Threshold (Yes/No)	If Implementation of Strategy Did Not Meet Threshold, Possible Reasons
National convenings	2 national convenings led by Pathway leadership team by the end of each year (SY 19/20, SY 20/21, SY 21/22)	2 national convenings held each year (SY 19/20, SY 20/21, SY 21/22)	Yes	N/A
Video recordings of UCIWP professional development	All core PD events video recorded and provided to other WP sites (SY 19/20, SY 20/21)	Videos of UCIWP SY 19/20 and 20/21 PD events covering 4 tutorials each year	Yes	N/A
Technology tools for classroom use	Instructional videos developed on 5 topics; sentence fluency tool; revision highlighter	Instructional videos and tools developed and available on Pathway website	Yes	N/A
Language Support Guidebook	Resource includes salient language features that teachers can use to support students' academic writing and is publicly available	Language Support Guidebook developed and available online	Yes	N/A
New research	2 journal publications and conference presentations	1 journal article and 1 conference presentation	Yes	N/A
Develop "teacher-consultant" (TC) classroom teachers to serve as Site PD Leads	Each expansion site (7) brings at least 1 TC to each national convening, beginning a year in advance of launching the PD with their partner district	TCs from 5 expansion sites attended at least 5 national convenings; TCs from 1 expansion site attended at least 4; TCs from 1 site attended at least 2 [Attendance data are missing for the national convening held in February 2021.]	No	One site joined the project too late for TCs to attend the national convenings the year before they launched PD with their partner district
	In the year prior to program launch or during Year 1 of implementation, at least 1 TC from each expansion site administers pretests and posttests and implements the cognitive strategies and revision tutorials	In 6 of the 7 sites, at least one TC partially implemented the core components of the Pathway model; no TC implemented the revision tutorial in its entirety	No	The full school year prior to the program launch in most sites (SY 19/20) was disrupted by the pandemic and no TCs were able to implement the revision tutorial in

Scale-Up Strategy	Threshold for Successful Implementation	Findings on Actual Level of Implementation	Implementation of Strategy Met or Exceeded Threshold (Yes/No)	If Implementation of Strategy Did Not Meet Threshold, Possible Reasons
				spring 2020. The subsequent year (SY 20/21), instruction was remote in most places, further challenging implementation
Develop Site Directors' (SDs') understanding of Pathway model	Each expansion SD attends at least one national convening per year	SDs from 5 expansion sites attended at least 5 national convenings; SD from 1 expansion site attended at least 4; SD from 1 site attended at least 2 [Attendance data are missing for the national convening held in February 2021.]	Yes	N/A
	In the year prior to program launch or during Year 1 of implementation, SD observes TCs' implementation of cognitive strategies and revision tutorials and/or debriefs video of classroom implementation with TCs	No SD observed TCs' implementation of cognitive strategies and revision tutorials and/or debriefed video of classroom implementation with TCs	No	As noted above, no TCs implemented all core components of the program in their own classrooms
Deploy "thinking partners" (TPs) to each expansion site	85% of the 7 UCI TPs attend at least 7 meetings per year with NWP	100% of TPs attended at least 7 meetings in SY 20/21; 6 of 7 TPs (86%) attended at least 7 meetings in SY 21/22.	Yes	N/A
	100% of TPs conduct annual site visit to each expansion site	Information not readily available	Information not readily available	N/A

Note. SD = Site Director; SY = school year; TC = teacher consultant; TP = thinking partner; UCIWP = UC Irvine Writing Project; PD = professional development; WP = Writing Project; NWP = National Writing Project.

Cost-Effectiveness Study

In this section, we describe the costs of this Pathway intervention relative to the number of students who benefit.

For costs, we took the grant costs (including nonfederal funds, or "matching" dollars) in each year students were served: 2019–20, 2020–21, and 2021–22. We subtracted the evaluation costs in each of these years (Exhibit 28).

Exhibit 28. Cost by Year

Year	Costs in Dollars
Year 2 (2019–20)	$1,443,895
Year 3 (2020–21)	$1,995,913
Year 4 (2021–22)	$2,220,241
Total	$5,660,049

Note. For costs, we used the costs as reported by the grantee as part of the required Annual Performance Reporting. Given the high costs of grant management not typically associated with program implementation (e.g., recruitment), we only include costs in years that students were served by the program.

We caution that this is a coarse estimate of costs. It includes out-of-pocket expenses that may not be relevant to school districts considering implementation (e.g., recruitment and grant administration that took place during these program years), and it does not include costs internal to school districts that may not have been represented in the grant costs (e.g., administrator time to support professional development). In addition, the estimated costs are average costs across years over 3 years of program implementation; the costs may differ for different years of implementation or a different duration of program adoption. Further, as with all estimates in this study, the findings may not be generalizable to studies in non-pandemic years.

For students served, we provide both a lower and upper bound (Exhibit 29). For the lower bound, we provide the unique count of students who were in a Pathway (i.e., focal) classroom for at least one full school year. This number should represent the students who received the full treatment, as described in the fidelity of implementation section.

We also provide the upper bound, which includes all students who were in a Pathway teacher's classroom during any year of the study. We include these students as teachers frequently reported using Pathway in non-focal classes. Further, at least one prior study of a similarly structured program (i.e., one that provided both instructional materials and professional development focused on teacher implementation of those materials) found that students are likely to benefit from the program even when teachers did not use the materials in the classroom (Arshan et al., 2024). Of course, teacher may also use these instructional strategies and materials in subsequent schools year, potentially impacting many more students.

Exhibit 29. Upper and Lower Bounds of Students Served and Associated Cost per Student

Assumption	Total Cost in Dollars	Students Served	Cost per Student
Conservative cost estimate (lower bound of students served)	$5,660,049	10,032	$564.20
Liberal cost estimate (upper bound of students served)	$5,660,049	19,198	$294.84

We provide these two bounded estimates of cost per student served, along with the standardized impact estimates for our three confirmatory contrasts (Exhibit 30).

Exhibit 30. Student Impacts and Cost-Per-Student

Student Outcome Measure	Impact Estimate (standardized effect size)	Liberal Cost-Per-Student	Conservative Cost-Per-Student
Holistic	.12	$294.84	$564.20
Conventions	.10	$294.84	$564.20
Word Count	.19	$294.84	$564.20

Conclusion

Although writing has become central to college- and career-ready standards within the past 15 years, evidence suggests writing remains a "neglected skill" in America's schools (Graham, 2019), and the pandemic may have exacerbated existing opportunity gaps across subjects (Fahle et al., 2023). UCIWP's Pathway program provides secondary ELA and ELD teachers with professional development and instructional materials to help them teach writing and support students in meeting state-adopted ELA standards and graduating from high school prepared for college and work.

With funding from a federal EIR grant, UCIWP extended capacity to deliver Pathway's teacher professional development in new contexts and validate the efficacy of the approach in these new settings. As a result of pandemic-related disruptions and adaptations, Pathway implementation looked different than expected. Teachers participated in fewer total hours of professional development than planned, and fewer teachers than expected were able to implement all three tutorials in their classrooms.

Despite these interruptions, the study found that Pathway led to positive, statistically significant impacts on students' writing. Thus, this evaluation—the second independent evaluation of Pathway when replicated by National Writing Project sites other than UCIWP—provides further evidence of Pathway's effectiveness and contributes to the research on writing instructional practices, teacher professional development, and scaling interventions.

References

Allen, J. P., Gregory, A., Mikami, A., Lun, J., Hamre, B., & Pianta, R. (2013). Observations of effective teacher-student interactions in secondary school classrooms: Predicting student achievement with the classroom assessment scoring system-secondary. *School Psychology Review*, *42*(1), 76–98. https://doi.org/10.1080/02796015.2013.12087492

Arshan, N., & Friedrich, L. (2017, March 1–4). *Development of an outcome measure to meet evaluation needs: The case of the analytic writing continuum for literary analysis* [Paper presentation]. Society for Research on Educational Effectiveness Spring Conference, Washington, DC, United States.

Arshan, N. L., Park, C. J., & Goetz, R. (2024). Supporting teachers in argument writing instruction at scale: A replication study of the College, Career, and Community Writers Program (C3WP). *Journal of Research on Educational Effectiveness*, *17*(1), 158–183. https://doi.org/10.1080/19345747.2023.2179953

Bacha, N. (2001). Writing evaluation: What can analytic versus holistic essay scoring tell us? *System*, *29*(3), 371–383. https://doi.org/10.1016/S0346-251X(01)00025-2

Bang, H. J. (2013). Reliability of National Writing Project's Analytic Writing Continuum Assessment System. *The Journal of Writing Assessment*, *6*(1). https://escholarship.org/uc/item/03g148gh

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to improve: How America's schools can get better at getting better*. Harvard Education Press.

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

Culham, R. (2003). *6+1 Traits of writing*. Scholastic Professional Books.

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute. https://learningpolicyinstitute.org/product/effective-teacher-professional-development-brief

Desimone, L. M., Porter, A. C., Garet, M., Yoon, K. S., & Birman, B. (2002). Does professional development change teachers' instruction? Results from a three-year study. *Educational Evaluation and Policy Analysis*, 24(2), 81–112. https://doi.org/10.3102/0013189X08331140

Fahle, E. M., Kane, T. J., Patterson, T., Reardon, S. F., Staiger, D. O., & Stuart, E. A. (2023). *School district and community factors associated with learning loss during the COVID- 19 pandemic*. Center for Education Policy Research at Harvard University. https://cepr.harvard.edu/sites/hwpi.harvard.edu/files/cepr/files/explaining_covid_losses_5.23.pdf

Garet, M. S., Heppen, J., Walters, K., Smith, T., & Yang, R. (2016). *Does content-focused teacher professional development work? Findings from three Institute of Education Sciences studies* (NCEE 2017-4010). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/pubs/20174010/

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

Graham, S. (2019). Changing how writing is taught. *Review of Research in Education*, *43*(1), 277–303. https://doi.org/10.3102/0091732X18821125

Graham, S., Bruch, J., Fitzgerald, J., Friedrich, L., Furgeson, J., Greene, K., Kim, J., Lyskawa, J., Olson, C. B., & Smither Wulsin, C. (2019). *Teaching secondary students to write effectively* (NCEE 2017-4002, Revised). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, What Works Clearinghouse. https://ies.ed.gov/ncee/WWC/PracticeGuide/22

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

Hunter, D. M., Jones, R. M., & Randhawa, B. S. (1996). The use of holistic versus analytic scoring for large-scale assessment of writing. *Canadian Journal of Program Evaluation*, *11*(2), 61–86. https://doi.org/10.3138/cjpe.11.003

Kane, M. T. (2013). Validation as a pragmatic, scientific activity. *Journal of Educational Measurement*, *50*(1), 115–122. https://doi.org/10.1111/jedm.12007

Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, *86*(4), 945–980. https://doi.org/10.3102/0034654315626800

Kim, J. S., Olson, C. B., Scarcella, R., Kramer, J., Pearson, M., van Dyk, D., Collins, P., & Land, R. E. (2011). A randomized experiment of a cognitive strategies approach to text-based analytical writing for mainstreamed Latino English language learners in grades 6 to 12. *Journal of Research on Educational Effectiveness*, *4*(3), 231–263. https://doi.org/10.1080/19345747.2010.523513

Lieberman, A., & Wood, D. R. (2002). From network learning to classroom teaching. *Journal of Educational Change*, *3*(3), 315–337. https://doi.org/10.1023/A:1021286014650

May, H., Perez-Johnson, I., Haimson, J., Sattar, S., & Gleason, P. (2009). *Using state tests in education experiments: A discussion of the issues* (NCEE 2009-013). 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=NCEE2009013

McCaffrey, D. F., Yuan, K., Savitsky, T. D., Lockwood, J. R., & Edelen, M. O. (2015). Uncovering multivariate structure in classroom observations in the presence of rater errors. *Educational Measurement: Issues and Practice*, *34*(2), 34–46. https://doi.org/10.1111/emip.12061

McDonald, S.-K., Keesler, V. A., Kauffman, N. J., & Schneider, B. (2006). Scaling-up exemplary interventions. *Educational Researcher*, *35*(3), 15–24. https://doi.org/10.3102/0013189X035003015

Morel, R. P., Coburn, C., Catterson, A. K., & Higgs, J. (2019). The multiple meanings of scale: Implications for researchers and practitioners. *Educational Researcher*, *48*(6), 369–377. https://doi.org/10.3102/0013189X19860531

National Center for Education Statistics. (2012). *The Nation's Report Card: Writing 2011* (NCES 2012-470). U.S. Department of Education, Institute of Education Sciences. https://nces.ed.gov/nationsreportcard/pdf/main2011/2012470.pdf

National Commission on Writing in America's Schools and Colleges. (2003). *The neglected "R": The need for a writing revolution*. College Entrance Examination Board. https://eric.ed.gov/?id=ED475856

Olson, C. B., Kim, J. S., Scarcella, R., Kramer, J., Pearson, M., van Dyk, D. A., & Land, R. E. (2012). Enhancing the interpretive reading and analytical writing of mainstreamed English learners in secondary school: Results from a randomized field trial using a cognitive strategies approach. *American Educational Research Journal*, *49*(2), 323–355. https://doi.org/10.3102/0002831212439434

Olson, C. B., & Land, R. (2007). A cognitive strategies approach to reading and writing instruction for English language learners in secondary school. *Research in the Teaching of English*, *41*(3), 269–303. https://doi.org/10.58680/rte20076014

Olson, C. B., & Land, R. (2008). Taking a reading/writing intervention for secondary English language learners on the road: Lessons learned from the Pathway Project. *Research in the Teaching of English*, *42*(3), 259–269. https://doi.org/10.58680/rte20086495

Olson, C. B., Matuchniak, T., Chung, H. Q., Stumpf, R., & Farkas, G. (2017). Reducing achievement gaps in academic writing for Latinos and English learners in grades 7–12. *Journal of Educational Psychology*, *109*(1), 1–21. https://doi.org/10.1037/edu0000095

Olson, C. B., Woodworth, K., Arshan, N., Black, R., Chung, H. Q., D'Aoust, C., Dewar, T., Friedrich, L., Godfrey, L., Land, R., Matuchniak, T., Scarcella, R., & Stowell, L. (2020). The pathway to academic success: Scaling up a text-based analytical writing intervention for Latinos and English learners in secondary school. *Journal of Educational Psychology*, *112*(4), 701–717. https://doi.org/10.1037/edu0000387

Slavin, R. E., Lake, C., Inns, A., Baye, A., Dachet, D., & Haslam, J. (2019). *A quantitative synthesis of research on writing approaches in grades 2 to 12*. Johns Hopkins University, Center for Research and Reform in Education. https://bestevidence.org/wp-content/uploads/2021/02/writing-april-2019.pdf

What Works Clearinghouse. (2021). *Study review protocol* (Version 1.0). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/WWC_Study_Review_Protocol_1_0.pdf

Wolf, B., & Harbatkin, E. (2023). Making sense of effect sizes: Systematic differences in intervention effect sizes by outcome measure type. *Journal of Research on Educational Effectiveness*, *16*(1), 134–161. https://doi.org/10.1080/19345747.2022.2071364

Woodworth, K. R., Arshan, N. L., & Gallagher, H. A. (2017). *UC Irvine Writing Project's Pathway to Academic Success program: An Investing in Innovation (i3) Validation grant evaluation* [Technical report]. SRI International. https://www.sri.com/wp-content/uploads/2021/12/pathway_i3_sri_technical_report_21dec17_final_in_jan.pdf

Appendix

Exhibit A-1. Number of Schools and Students in the Randomized Sample and With Writing Essays

Outcome Measure	Comparison Group				Treatment Group			
	Clusters		Students		Clusters		Students	
	# Randomized	# With Essay	# Randomized	# With Essay	# Randomized	# With Essay	# Randomized	# With Essay
Holistic	23	23	3,551	1,495	23	23	3,856	1,726
Conventions	23	23	3,551	1,495	23	23	3,856	1,726
Word count	23	23	3,551	1,495	23	23	3,856	1,726

Note. The overall student attrition rate for the 46 study schools was 57%, with a differential attrition rate of 2.7%, which exceeds What Works Clearinghouse's optimistic boundary of 2.1%. The relatively high attrition rates prevent an analysis of only Year 1 students from providing the highest validity from the randomized controlled trial design. As a result, we did not pursue a separate Year 1 impact analysis for students at the time of randomization.

Exhibit A-2. Control Variables

Level	Control Variable
School	Percent of students eligible for free or reduced-price lunch (FRPL) Percent English learner students Percent African American students, percent Hispanic students Randomization pair blocking indicators
Student	Baseline writing score (same as outcome) Gender – female Race – Asian, Black, Hispanic, White, Other Receives free or reduced-price lunch (FRPL status) Limited English Proficiency (LEP) Individualized Education Program (IEP) Grade level Outcome year indicators (2020–21 and 2021–22 vs. 2019–20) Student group indicators (Year 2 and both-year vs. Year 1) Interaction of baseline score for both year status

Exhibit A-3. Baseline Equivalence for Additional Writing Attributes, Student Writing Outcomes Sample

Attribute	Comparison Group			Treatment Group				
	Sample Size	Mean	*SD*	Sample Size	Mean	*SD*	Treatment–Comparison Difference	Standardized Difference
Content	1,904	2.87	1.25	2,232	2.69	1.26	-0.11	-0.09
Structure	1,904	2.82	1.24	2,232	2.63	1.24	-0.13	-0.11
Stance	1,904	2.88	1.27	2,232	2.69	1.28	-0.14	-0.11

Note. Exhibit presents sample sizes, means, and standard deviations of baseline measures for additional Analytic Writing Continuum for Literary Analysis (AWC-LA) writing outcomes for comparison and treatment group separately, the adjusted treatment-comparison differences from the model, and standardized differences (Hedge's *g*) based on the pooled standard deviation of treatment and comparison groups. All standardized differences in outcomes are smaller than 0.25, establishing baseline equivalence for this set of analysis. *SD* = standard deviation.

Exhibit A-4. Descriptives of Outcome Scores and Estimated Program Impact on Additional AWC-LA Writing Attributes

Attribute	Comparison Group			Treatment Group						
	Sample Size	Mean	*SD*	Sample Size	Model-Adjusted Mean	*SD*	Treatment –Comp. Difference	*SE*	Standardized Difference	*p*-Value
Content	1,904	3.35	1.34	2,232	3.50	1.39	0.15	0.06	0.12	.011
Structure	1,904	3.27	1.31	2,232	3.45	1.37	0.18	0.06	0.14	.002
Stance	1,904	3.36	1.36	2,232	3.51	1.40	0.15	0.06	0.12	.015

Note. Exhibit has descriptive information and impact estimates for additional Analytic Writing Continuum for Literary Analysis (AWC-LA) writing outcomes. For content, structure, and stance, the Pathway project increased the performance of students in the treatment group 0.11, 0.13, and 0.11 standard deviations more than that of students in the comparison group ($p < .05$) on the content, structure, and stance scores, respectively. *SD* = standard deviation; *SE* = standard error.

SRI Education, a division of SRI, is helping federal and state agencies, school districts, major foundations, nonprofit organizations, and international and commercial clients tackle some of the most complex issues in education to help students succeed. Our mission is ***to reduce barriers, optimize outcomes, and ensure educational equity for all children, youth, and families***. We do this by conducting high-quality research, supporting use of data and evidence, helping to strengthen state and local systems, and developing tools that improve teaching and accelerate and deepen learning. Our work covers a range of topics: early learning and development, disability and inclusion, supporting multilingual learners, student behavior and well-being, teaching quality, digital learning, STEM and computer science, literacy and language arts, and college and career pathways. **We believe diversity in our organization and project teams leads to better and more equitable research and technical assistance, resulting in improved outcomes for all.**

Silicon Valley
(SRI headquarters)
333 Ravenswood Avenue
Menlo Park, CA 94025
1.650.859.2000

Washington, DC
1100 Wilson Boulevard
Suite 2800
Arlington, VA 22209
1.703.524.2053

education@sri.com

www.sri.com/education-learning/