Title

The Pathway to Academic Success: Scaling Up a Text-Based Analytical Writing Intervention for Latinos and English Learners in Secondary School

Journal

Journal of Educational Psychology, 112(4)

ISSN

0022-0663

Authors

Olson, Carol Booth
Woodworth, Katrina
Arshan, Nicole
et al.

Publication Date

2020

DOI

10.1037/edu0000387

Peer reviewed

Journal of Educational Psychology

The Pathway to Academic Success: Scaling Up a Text-Based Analytical Writing Intervention for Latinos and English Learners in Secondary School

--Manuscript Draft--

Manuscript Number:	EDU-2018-3232R2
Full Title:	The Pathway to Academic Success: Scaling Up a Text-Based Analytical Writing Intervention for Latinos and English Learners in Secondary School
Abstract:	This study reports findings from a multisite cluster randomized controlled trial designed to validate and scale up an existing successful professional development program that uses a cognitive strategies approach to text-based analytical writing. The Pathway to Academic Success Project worked with partner districts affiliated with four National Writing Project (NWP) sites in southern California. Informed by a wide body of research on the efficacy of strategy instruction to enhance students' academic literacy, the intervention aimed to help secondary school students, particularly Latinos and mainstreamed English learners, to develop the academic writing skills called for in the rigorous Common Core State Standards for English Language Arts. Two hundred and thirty teachers from partner districts affiliated with the NWP sites were stratified by school and grade and then randomly assigned to the treatment or control group. Treatment teachers participated in 46 hours of training and learned how to apply cognitive strategies by using an on-demand writing assessment to help students understand, interpret and write analytical essays about nonfiction texts. Multilevel models revealed significant effects on a holistic measure of an on-demand writing assessment (d =.32) as well as on four analytic attributes: content (d =.31); structure (d =.29); fluency (d =.27); and conventions (d =.32). Four dimensions of scaling up—spread, reform ownership, depth, and sustainability— are also discussed.
Article Type:	Article
Keywords:	adolescents; Scaling Up; Second Language Learning; Teacher Professional Development; writing
Corresponding Author:	Carol Booth Olson University of California, Irvine Irvine, California UNITED STATES
Corresponding Author E-Mail:	cbolson@uci.edu
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of California, Irvine
Other Authors:	Katrina Woodworth
	Nicole Arshan
	Rebecca Black
	Huy Q. Chung
	Catherine D'Aoust
	Tim Dewar
	Linda Friedrich
	Lauren Godfrey
	Robert Land
	Tina Matuchniak
	Robin Scarcella
	Laurie Stowell
Corresponding Author's Secondary	

Institution:	
First Author:	Carol Booth Olson
Order of Authors Secondary Information:	
Manuscript Region of Origin:	UNITED STATES
Suggested Reviewers:	Judith Langer University at Albany State University of New York jlanger@alabany.edu Dr. Langer is an expert in interventions for English Language Arts
Opposed Reviewers:	
Response to Reviewers:	
Order of Authors:	Carol Booth Olson
	Katrina Woodworth
	Nicole Arshan
	Rebecca Black
	Huy Q. Chung
	Catherine D'Aoust
	Tim Dewar
	Linda Friedrich
	Lauren Godfrey
	Robert Land
	Tina Matuchniak
	Robin Scarcella
	Laurie Stowell
Manuscript Classifications:	Minority Students; Teacher Effects; Writing

Authors' Responses

We have taken your suggestions into account when revising our manuscript for a second time. Per our communication with you, we have addressed your concerns with the previous draft.

Additionally, per author submission guidance we would like to ask your permission to add supplementary materials with online access as the inclusion of these tables and figures would make the manuscript longer than necessary, but are important for readers to understand the full scope of the project and project's findings.

We hope these revisions have made for a stronger manuscript. Thank you for your time and consideration.

Cover Letter

UNIVERSITY OF CALIFORNIA, IRVINE

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ

School of Education
www.gse.uci.edu

UCI Writing Project
401 East Peltason, Suite 3100
Irvine, CA 92697-5510
(949) 824-5922
(949) 824-6685 FAX

May 15, 2019

Dr. Steve Graham
Editor
Journal of Educational Psychology

Dear Dr. Graham,

Thank you for providing critical suggestions on our manuscript, The Pathway to Academic Success: Scaling Up a Text-Based Analytical Writing Intervention for Latinos and English Learners in Secondary School.

We have taken your suggestions into account when revising our manuscript for a second time. Per our communication with you, we have addressed your concerns with the previous draft.

Additionally, per author submission guidance we would like to ask your permission to add supplementary materials with online access as the inclusion of these tables and figures would make the manuscript longer than necessary, but are important for readers to understand the full scope of the project and project's findings.

We hope these revisions have made for a stronger manuscript. Thank you for your time and consideration.

Sincerely,

Carol Booth Olson

Carol Booth Olson, Ph.D.
Professor of Education
School of Education
University of California, Irvine

Title page with All Author Information

The Pathway to Academic Success: Scaling Up a Text-Based Analytical Writing Intervention for Latinos and English Learners in Secondary School

Carol Booth Olson, (corresponding author)
Professor of Education, Director of the UCI Writing Project
School of Education
University of California, Irvine
(949) 824-7842
cbolson@uci.edu

Katrina Woodworth
SRI, International

Nicole Arshan
SRI, International

Rebecca Black
University of California, Irvine

Huy Q. Chung
University of California, Irvine

Catherine D'Aoust
University of California, Irvine

Tim Dewar
University of California, Santa Barbara

Linda Friedrich
National Writing Project

Lauren Godfrey
University of California, Irvine

Robert Land
California State University, Los Angeles

Tina Matuchniak
California State University, Long Beach

Robin Scarcella
University of California, Irvine

Laurie Stowell
California State University, San Marcos

Abstract

This study reports findings from a multisite cluster randomized controlled trial designed to validate and scale up an existing successful professional development program that uses a cognitive strategies approach to text-based analytical writing. The Pathway to Academic Success Project worked with partner districts affiliated with four National Writing Project (NWP) sites in southern California. Informed by a wide body of research on the efficacy of strategy instruction to enhance students' academic literacy, the intervention aimed to help secondary school students, particularly Latinos and mainstreamed English learners, to develop the academic writing skills called for in the rigorous Common Core State Standards for English Language Arts. Two hundred and thirty teachers from partner districts affiliated with the NWP sites were stratified by school and grade and then randomly assigned to the treatment or control group. Treatment teachers participated in 46 hours of training and learned how to apply cognitive strategies by using an on-demand writing assessment to help students understand, interpret and write analytical essays about nonfiction texts. Multilevel models revealed significant effects on a holistic measure of an on-demand writing assessment (*d* =.32) as well as on four analytic attributes: content (*d* =.31); structure (*d* =.29); fluency (*d* =.27); and conventions (*d* =.32). Four dimensions of scaling up—spread, reform ownership, depth, and sustainability— are also discussed.

Key words: *Adolescents; Scaling Up; Second Language Learning; Teacher Professional Development; Writing*

Educational Impact and Implications Statement

This study sought to address and validate solutions to persistent educational challenges by scaling up an intervention that trains English language arts teachers to take a cognitive strategies approach to literacy instruction to prepare Latinos and mainstreamed English learners (ELs) in high need, low SES schools to meet the state-adopted English language arts standards and improve their text-based analytical writing. Results revealed that students in the classes of treatment teachers significantly outgained their peers in the control condition on an on-demand writing assessment and that treatment Latinos and ELs achieved parity with White and non-EL control students. Given the dearth of research regarding effective literacy instruction for Latinos and mainstreamed ELs at the secondary level, these results are important for practitioners, intervention developers, and policy makers. The study also sheds light on strategies for moving beyond the spread of an intervention to larger numbers of participants to genuinely cultivating a shift in reform ownership, generating a depth of understanding, and planning for sustainability among those delivering and replicating the intervention, and within districts, schools, and teachers.

The Pathway to Academic Success: Scaling Up a Text-Based Analytical Writing Intervention for Latinos and English Learners in Secondary School

Re-Submitted to the Journal of Educational Psychology

May 15, 2019

> "In principle, it might seem that scaling up would be a simple, even trivial, task: One simply takes an intervention that has been used successfully on a small scale and applies it on a large scale. In practice, however, there are many difficulties associated with scaling up."
>
> Sternberg, Birney, Kirlik, Stemler, Jarvin, & Grigorenko, 2006 (p. 206)

Increasingly, government agencies, policy makers, researchers, and educators have turned their attention to expanding the reach of evidence-based interventions by scaling them up to new and larger contexts. However, as Sternberg and colleagues remind us, successfully transporting an intervention that has been developed and tailored to meet the needs of teachers and students in one location and spreading reform efforts to multiple districts in diverse geographical areas is not as easy as it sounds. In fact, in her seminal article on rethinking scale (2003), Coburn has identified the issue of scale as "one of the key challenges of educational reform" (p. 3).

In 2014, a site of the National Writing Project at the University of California, Irvine obtained a four-year Validation grant from the U.S. Department of Education, Investment in Innovation (i3) to scale up an intervention which had been developed and successfully implemented in several districts in Orange County, California to three other National Writing Project (NWP) sites at three other universities in the southern California region. The focus of the Pathway to Academic Success Project was to enhance the academic literacy of mainstreamed, predominately Latino English learners (ELs) and the native English speaking peers in their grades 7-12 classrooms in largely urban, high need schools with large percentages of students receiving free and reduced-price lunch. The goal was to help students receiving the treatment to meet California Common Core State

Standards for English Language Arts (CCSS-ELA) in order to graduate from high school and become college bound. Specifically, the intervention provided ongoing, sustained evidence-based professional development (PD) for teachers of secondary ELs focused on how to explicitly teach, model, and scaffold guided instruction in the cognitive strategies (or thinking tools) that research indicates experienced readers and writers access when they construct meaning from and with texts. The aim was to prepare these EL teachers' students to become strategic readers able to analyze and interpret complex texts and analytical writers capable of producing well-reasoned essays supported with textual evidence. Each of the NWP Site Directors selected a high need school district or districts to serve as its PD partner or partners. (See the list of Site Directors and partner districts included as Figure A in the online supplementary materials).

Coburn (2003) argues that expanding a reform to multiple settings is a "necessary but insufficient condition for scale" (p.4). Scaling up does, indeed, involve an increase in users, or what Coburn calls *spread*, but engendering a shift in *reform ownership* from the initial developers of an intervention to other facilitators and from those facilitators to a growing number of districts, schools, and teachers, as well as effecting a *depth* of change in teachers' beliefs and practices, and promoting the *sustainability* of intervention effects after initial implementation are also dimensions necessary to successfully bring an intervention to scale. In this article, we report on Year 1 findings from the randomized controlled field trial designed to test whether a successful intervention created and previously delivered by the program developer could be successfully scaled up to new locations and effectively delivered by literacy specialists not previously affiliated with

the intervention to improve writing outcomes for all students, but particularly to reduce achievement gaps in academic writing for Latinos and ELs in secondary school.

Targeting a Text-Based Analytical Writing Intervention for Latinos and English Learners

In its vision of what it means to be literate in the twenty-first century, the CCSS-ELA prioritize the ability to "read closely to determine what the text says explicitly and to make logical inferences from it," (p.10) and to "write arguments to support claims in an analysis of substantive topics or texts, using valid reasoning and relevant and sufficient evidence" (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010, p.18). As is evident from these anchor standards, the CCSS-ELA, and other state standards developed and adopted since 2010, set a high bar for all students and place a premium on text-based analytical writing, identifying the ability to analyze and interpret challenging texts and to write about those texts using academic discourse in extended pieces of writing as the key to academic success. However, results from the most recent administration of the National Assessment of Educational Progress (NAEP) in writing in 2011 (U.S. Department of Education, Institute of Educational Sciences, National Center for Education Statistics, 2012) and the administration of the NAEP in reading in 2015 (U.S. Department of Education, Institute of Educational Sciences, National Center for Education Statistics, 2016) indicate that today's secondary students face considerable challenges in meeting these standards. Nationally, only 20% of 8th graders and 18% of 12th graders scored at proficient or above in writing; additionally only 27% of both 8th and 12th graders,

respectively, scored at proficient or above in reading. As Table 1 indicates, large disparities exist between the performance of White, Hispanic, and Black students:

[Insert Table 1 here.]

Currently, Latinos are the largest ethnic group in California. One out of every two California youth under the age of 25 is Latino (United States Consensus Bureau, 2015), and two-thirds of the state's Latino youth are concentrated in southern California (United States Census Bureau, 2017). Today, among all ethnic groups in California, Latinos are the least likely to obtain a college degree (National Center for Education Statistics, 2017). Hence, in a recent report on supporting the educational success of Latino students in California, the Education Trust-West concludes, "Unless we accelerate the pace of change, our state is sending a disturbing message about the contributions, work, and potential of Latino students" (p. 7) Given that by 2020, one in four children enrolled in America's K-12 public schools will be Latino (Maxwell, 2012), improving educational outcomes for Latinos is clearly also a pressing national concern.

While the Latino population continues to rise in California and nationally, ELs represent the fastest growing segment of the K-12 population with the largest increases occurring in grade 7-12 (U.S. Department of Education, Institute of Educational Sciences, National Center for Education Statistics, 2017). California leads the nation with almost 23% ELs, but many other states such as New York, Florida, Colorado, Washington, Nevada, Texas, Arizona, Illinois, and Minnesota serve large percentages of ELs (U.S. Department of Education, Institute of Educational Sciences, National Center for Education Statistics, 2016). Although ELs in the United States speak more than 350 languages, 73% speak Spanish as their first language (Batalova & McHugh, 2010), 40%

have origins in Mexico (Hernandez, Denton, & Macartney, 2008), and 60% of ELs in grades 6 through 12 come from low-income families (Batalova, Fix, & Murray, 2005; Capps et al., 2005). The largest numbers of ELs in our schools today are referred to as long-term ELs (LTELs) (Menken & Kleyn, 2009). These are students who have been educated in the United States since age six, are doing poorly in school, have major gaps in knowledge because their schooling has often been disrupted, are failing to acquire academic language, and are struggling to do well in high school (Olsen, 2010). Limited in their knowledge of academic registers in any language, these students are often mainstreamed into regular English language arts classrooms.

The most alarming result from the NAEP assessments is that only 1% of ELs at both grades 8 and 12 scored at proficient or above in writing and only 4% of ELs scored at proficient or above in reading (U.S. Department of Education, National Center for Education Statistics 2012 & 2016). This disparity in performance is likely to be a major hindrance for ELs' college access and persistence since academic preparation in high school is a major predictor of college success (Adelman, 2006; Kanno & Cromley, 2015). Not only are ELs less likely than their non-EL peers to graduate from high school, they also enroll in college and graduate from college at far lower rates (Nuñez, Rios-Aguilar, Kanno, & Flores, 2016), and studies have shown that academic writing is "the linguistic challenge that plagues" EL college students the most (Kanno & Cromley, 2015). Because text-based analytical writing is a gatekeeper for college access and persistence and a "threshold skill" for hiring and promotion for salaried workers (National Commission on Writing for America's Families, Schools, and Colleges, 2004), failure to close these

achievement gaps in academic writing will have serious social and economic consequences.

The paucity of empirical studies on effective practices for teaching text-based analytical writing to ELs reveals that this research area is nearly untapped (August & Shanahan, 2006; Fitzgerald, 2017). Most of the studies are small in size and qualitative in nature. Fitzgerald and Amendum (2007), for example, report no empirical studies of grade 6-12 writing instruction in their meta-analysis that involved 1988-2003 research studies of the K-12 writing instruction for ELs in the United States, and emphasize the growing need to investigate effective practices for teaching writing to adolescent ELs. This lack of research leaves the teachers of over 4.8 million EL students largely to speculate about how best to teach their students. How can they teach them to meet rigorous standards when they have so little training in how to diversify instruction in order to meet the needs of ELs, especially given that writing instruction is one of the least addressed areas in teacher preparation (Kiuhara, Graham, & Hawken, 2009)?

Taking A Cognitive Strategies Approach to Teaching Text-Based Analytical Writing to Latinos and Mainstreamed ELs: A Review of the Relevant Scholarship

The Pathway intervention takes a cognitive strategies approach to closing the achievement gap between ELs and their native English-speaking peers in the area of text-based analytical writing. Cognitive strategies are acts of mind, or thinking tools, such as planning and goal setting tapping prior knowledge, making connections, monitoring, forming interpretations, reflecting and relating, evaluation, etc., that research indicates readers and writers use to construct meaning (Olson, 2011). There is widespread agreement in the research community that cognitive strategy use enhances reading

comprehension (Block & Pressley, 2002; National Institute of Child Health and Human Development, 2000; Tierney & Pearson, 1983). Similarly, Graham and Perin (2007) indicate that strategy instruction is the most effective of eleven key elements of writing instruction (d=.82) for all students and particularly for students who find writing challenging. In fact, in the WWC Practice Guide *Teaching Secondary Students to Write Effectively* (Graham et al., 2016), the expert panel determined as its number one recommendation, with the highest level of evidence, that teachers should explicitly teach, model, and enable students to practice and reflect upon writing strategies and concluded, "teaching students cognitive strategies is one way to develop their strategic thinking skills, ultimately helping them to write more effectively" (p. 7). As MacArthur (2019) points out, one of the most extensively studied and effective approaches to teaching writing is cognitive strategy instruction. Numerous studies have been conducted across the grade levels to establish the efficacy of strategy instruction in writing, including those by Englert, Raphael, Anderson, Anthony, and Stevens (1991) at the elementary level, De La Paz and Graham (2002) at the secondary level, and MacArthur, Philippakos, and Ianetta (2015) with college basic writers, among others.

Although Graham and Perin (2007) note the dearth of experimental studies conducted with low achieving writers from low income families in inner city settings, and especially with low English language proficiency, evidence also exists of the positive impact of cognitive strategy instruction on the literacy of ELs. Short and Fitzsimmons (2007) hypothesize that strategy instruction is especially effective for ELs because it provides them with an explicit focus on language, increases their exposure to academic texts, makes the texts they read comprehensible, gives them multiple opportunities to

affirm or correct their understanding and use of language, assists them in retrieving new language features and in using these features for academic purposes, and provides them with the means of learning language on their own, outside of class. They further hypothesize that mainstreamed adolescent ELs with an intermediate level of English proficiency, who represent the majority of Long Term English Learners (LTELs) in many states (Olsen, 2010), have sufficient proficiency to benefit from strategy instruction (Echevarria, Short & Vogt, 2008; Short & Fitzsimmons, 2007). Explicitly teaching strategic reading and writing behaviors to ELs can help them engage with complex texts and convey those interpretations in well-reasoned essays to meet the CCSS-ELA and other state-adopted standards (August & Shanahan, 2006; Bunch, Kibler & Pimentel 2012), enabling them to write compositions judged to be higher in quality and displaying more depth of interpretation, greater clarity of thesis, and better idea organization (Fitzgerald, 2017).

Background on the Pathway to Academic Success Project: Evidence from Previous Studies

The intent of the Pathway Project is to address and validate solutions to persistent educational challenges by providing teachers with curriculum materials and instructional practices to prepare Latinos and mainstreamed ELs in high-need schools to successfully complete courses in core academic subjects, and to meet their state-adopted English language arts standards, in order to graduate from high school and to become college-bound and career ready. The treatment is an intensive 46-hour PD program (via six full-day released days interspersed throughout the school year and five 2-hour after-school sessions) in which secondary teachers learn how to integrate cognitive strategy

instruction into process writing to improve students', specifically high-need students and mainstreamed ELs', interpretive reading and text-based analytical writing. This is done by (a) using a cognitive strategies approach to reading and writing instruction, (b) instructing students to revise a pretest on-demand writing assessment into multiple draft essays, and (c) receiving ongoing support from experienced intervention teachers who serve as coaches to teachers in the experimental condition. Pathway is designed as a two-year intervention, where the first year introduces the program to teachers and the second year refines and reinforces their work with it. This paper reports only on Year 1 of the program; the PD offered in Year 2 is similar.

The intervention has a twenty-year track record of producing evidence of project effectiveness and, with each iteration, the intervention continually evolved to meet the changing needs of teachers, students, and state and national literacy standards. Developed by the UCI Writing Project, the Pathway Project began with an 8 year quasi-experimental longitudinal study in Santa Ana Unified School District (SAUSD), a large, urban district (98% Latino, 84% Free and Reduced Price Lunch, 88% mainstreamed ELs) that yielded an average effect size of .34 across the eight years of implementation (Olson & Land, 2007). A second quasi-experiment, funded by the California Postsecondary Commission (CPEC), was launched in two other unified school districts, Paramount and Lynwood, that yielded effect sizes of .63 and .27 in Years 1 and 2, respectively (Olson & Land, 2008). The project then received Institute of Education Sciences (IES) funding to conduct a Goal 3 Efficacy cluster randomized controlled trial in SAUSD. Year 1 of that RCT (Kim et al., 2011) yielded an effect size of .35.Year 2 of the RCT yielded an effect size of .67 (Olson et al., 2012). In both years of the study, there were statistically significant

effects on the writing subtest of the California Standards Test *(d=.10).* Since the most rigorous research was conducted in SAUSD, the project applied for and received an Office of English Language Acquisition (OELA) grant to conduct an RCT to replicate the project in Anaheim Union High School District. This study also yielded significant and positive results (Olson et al. 2017--Year 1, *d*=.48*;* Year 2, *d*=.60). 10th grade ELs in the treatment group in Year 2 also passed the CAHSEE at 20 percentage points higher than the state pass rate (treatment = 57.9%; State = 38%). This strong record of statistically significant results positioned the project to compete well for an i3 Validation/Expansion award that funded the current scale-up study.

Research Questions

Because of both the size and scope of the intervention, and requirement for an independent evaluator to collect and analyze all impact data, SRI International, an independent, nonprofit research center, conducted the intervention i3 evaluation using a cluster randomized controlled trial. The primary research questions and hypotheses were as follows:

1. *What impact did the intervention have on teacher practice?*

Based on prior research results, we hypothesized that the intervention would increase teacher practice in reading and writing instruction, resulting in more time spent on higher level meaning making of complex texts, increased time on writing instruction, and more time and attention spent on revision.

2. *What processes or structures facilitated the replication and scaling up of the Pathway Project intervention?*

Since this was our first attempt at scaling up this intervention to a level that PD in three sites would not be delivered by the program developer, our hypotheses for this research question were tentative at best. But we surmised that building a relationship of trust among the Site Directors, providing high quality ready to-use tutorials and instructional materials, establishing non-negotiables, and leaving room for Site Directors to develop reform ownership and make modifications to suit their teachers' and students' needs would all facilitate replication and scaling up.

3. *What impact did assignment to the treatment have on a holistic measure of students' analytical writing and on component measures (i.e., content, structure, sentence complexity, conventions)?*

Based on prior research, we hypothesized that the treatment students would display a distinctly more developed sense of analytical writing and outgain students in the control condition on a holistic measure of writing. Since this was the first time we had used the NWP Analytic Writing Continuum which results in both holistic and analytic scores, we expected treatment students to outscore controls on content but were uncertain about what the results would be for structure, fluency, and conventions.

4. *Did the intervention have differential impacts on students' analytical writing for girls, English learners, or Latino students compared to boys, non-English learners, and White students?*

Based on prior analyses and studies, we hypothesized that the intervention would boost Latino and English learner students' analytical writing abilities.

5. *What impact did the intervention have on students' state standardized test scores in ELA (i.e., on the Smarter Balanced assessment)?*

Based on prior analyses and studies, we were cautiously optimistic that the intervention would have a small but statistically significant and positive effect size for treatment students. However, given that this was the first year the Smarter Balanced Test was administered in California and three sites were implementing the intervention for the first time, we were prepared that our study might yield a null finding in this area.

Theoretical Frame

The intervention is informed by cognitive, sociocognitive, and sociocultural theory. In their cognitive process theory of writing, Flower and Hayes (1981) posit that writing is best understood "as a set of distinct thinking processes which writers orchestrate and organize during the act of composing" (p. 275), including planning, organizing, goal setting, translating, monitoring, reviewing, evaluating, and revising. They liken these processes to a "writer's tool kit" (p. 285), which is not constrained by any fixed order or series of stages. Similarly, Tierney and Pearson (1983) propose that "reading and writing are essentially similar processes of meaning construction" (p. 568), and that readers compose meaning, creating drafts of their understanding, just like writers do. This concept that reading and writing are reciprocal processes of meaning construction links directly to the creation of the reader's and writer's cognitive strategy tool kit that serves as the centerpiece of the Pathway intervention.

In describing the difficulty of composing written texts, Flower and Hayes (1980) aptly describe writers as simultaneously juggling "a number of demands being made on conscious attention" (p. 32). While all learners face similar cognitive, linguistic, communicative, contextual, and textual constraints when learning to write (Frederiksen & Dominic, 1981), the difficulties younger, inexperienced, and underprepared students face

are magnified. For these students, juggling constraints can cause cognitive overload. For example, ELs are often cognitively overloaded, especially in mainstreamed classrooms where they are held to the same performance standards as native English speakers (Short & Fitzsimmons, 2007). They face the dual challenge of simultaneously learning how to write while they are still developing proficiency in the English language. Both challenges exert considerable and sometimes competing demands on the cognitive system.

Graham (2018) has pointed out that "available cognitive models mostly ignore cultural, social, political, and historical influences on writing development" (p. 272), asserting that writing is "inherently a social activity, situated within a specific context" (p. 273). This view echoes Langer (1991) who, drawing on Vygotsky (1986), suggests that literacy is the ability to think and reason like a literate person within a particular society. She notes, "as children learn to engage in literate behaviors to serve the functions and reach the ends they see modeled around then, they become literate in a culturally specific way; they use certain cognitive strategies to structure their thoughts and complete their tasks, and not others." (p. 17) From a sociocognitive perspective, teachers should pay more attention to the social purposes to which literacy skills are applied, and should go beyond delivering lessons on content to impart strategies for thinking necessary to complete literacy tasks, first with guidance and, ultimately, independently.

Finally, sociocultural theory views meaning as being "negotiated at the intersection of individuals, culture, and activity" (Englert, Mariage, & Dunsmore, 2006, p. 208). Three tenets of sociocultural theory are applicable to the intervention (Adapted from Englert, Mariage, & Dunsmore, 2006): 1) Sociognitive apprenticeships: Offering cognitive apprenticeships that support novices in participation and performance of a

discipline, including the acquisition of discourses, tools and actions; 2) Procedural facilitators and tools: Supporting cognitive performance in advance of independent performance through the provision of cultural tools and procedural facilitators such as diagrams, graphic organizers, writing symbols, text structures, mnemonics, etc., to help students produce well written texts; 3) Communities of practice: Establishing communities of practice that emphasize knowledge-construction and knowledge dissemination.

Drawing from cognitive, sociocognitive, and sociocultural theory, the intervention focuses on cognitive strategy use as a vehicle for higher level thinking, uses an apprentice model where the teacher serves as a senior member of a learning community, provides a wide array of mental, linguistic and physical tools to promote cognitive strategy use, and promotes collaboration among teachers, between teachers and students, and among students. The efficacy of the intervention also owes much to the teachers' teaching-teachers model of the National Writing Project (NWP), with its inherent respect for the capacity of practitioners to generate and use knowledge to inform and improve their practice. Ongoing and sustained PD based on the analysis of student work contributes to the establishment of a professional learning community dedicated to the academic progress of students. This collaboration creates a network of language arts classes in which highly trained teachers prepare all students, but specifically Latinos and ELs, to develop the interpretive reading and analytical writing abilities necessary for academic success. The NWP model has much in common with Graham's (2018) writer(s)-within-community model of writing, a basic tenet of which is that "writing cannot be fully understood without considering how the communities in which it takes place and those

involved in creating it evolve, including how community and individuals reciprocally influence each other" (p. 273). Additionally, the NWP teachers-teaching-teachers philosophy, embraced by the intervention, is a vehicle for creating the reform ownership espoused by Coburn (2003) as being essential to scaling up.

Theory of Action

The theory of action underlying the Pathway Project is that research-based guidance delivered by project partners via various cognitive-strategy-based Pathway curriculum materials is delivered as outputs over two-full years and implemented by treatment teachers. This results in increased instructional time on writing, teacher knowledge of evidence-based practices and expertise in teaching reading and writing strategies as well as more frequent explicit instruction in strategies for reading and writing, modeling, guided practice, text-based instruction, connections to prior knowledge and personal/cultural experience, and more frequent revisions of academic writing. A mid-term outcome includes increased student text-based writing ability. Ultimately, the long-term outcomes are increased English language arts test scores for all students, but especially high- need students, including ELs, and increased high-school graduation rates for those students (See the Logic Model included as Figure B in the online supplementary materials).

Method

The evaluation was a teacher-randomized controlled trial measuring the impact of the intervention on teachers' 7th- through 12th-grade English language arts (ELA) instructional practices and students' analytical writing achievement.

Inclusion and exclusion. SRI developed inclusion and exclusion criteria for NWP sites and local education agencies/districts, teachers, focal classes, and students to ensure that comparison sites were similar to each other as much as possible.

NWP sites. The study site recruited three additional NWP sites with (1) reach across the southern California region, (2) experienced leadership, and (3) large numbers of ELs in their service areas. The four Writing Project sites, in turn, recruited districts in their service areas with high proportions of ELs in grades 7-12 and a desire to implement the intervention. Although two of the sites had some experience working in the districts they recruited, none had worked on a cognitive strategies approach to writing with any of the participating districts.

Teacher sample. NWP sites worked with their partner districts to identify eligible teachers and target them for recruitment. To be eligible, teachers must have been the primary ELA/ELD teacher in at least one eligible ELA class in grades 7–12 and made a 2-year commitment to the study. In addition, teachers agreed to implement the intervention in their focal class and not to share intervention materials with control teachers if randomized into the treatment condition.

Focal class and student sample. Eligible focal classes were ELA and ELD classes in grades 7 through 12, excluding any classes for which the writing assessment would be inappropriate (i.e., self-contained special education classes and self-contained ELD classes serving introductory EL students).

Participant characteristics. The following sections highlight the key demographic makeup of our participating districts, teachers, and students.

District demographics. The four districts had similar demographics in terms of the total number of English learners served (M=36%) and total numbers of high need students served (M=62%). All districts served a majority of Hispanic students (55% to 80%), followed by White students (4% to 30%), then Asian students (3% to 12%), and then Black students (2% to 3%), and the rest are students of other ethnicities.

Teacher demographics. The teacher sample consisted of 230 secondary ELA and ELD teachers across grades 7–12, from 40 schools; 113 teachers were assigned to treatment and 117 to control. No in-mover teachers were permitted into the study sample. On average, teachers in the study were similar across treatment conditions in both their overall number of years teaching experience (M = 12.21, SD = 7.17), $t(227) = -.78$, p = .43, and the proportion that were first-year teachers, (M = 6%), $\chi^2(1, N = 229) = 1.99$, p = .16. Teacher attrition from the student impact analyses was 2% at the end of Year 1 (1% for treatment and 3% for control).

Student demographics. Students in both treatment and control conditions were roughly equivalent across most demographic characteristics (e.g., English learner status and race/ethnicity). (See Table A in the online supplementary materials).

Sampling procedures. To achieve the desired analytic sample size, SRI added to the assigned sample until they had four students per classroom in the analytic sample (i.e., with both pretest and posttest student writing). The students in the assigned sample without posttest writing were considered attrition:

$$Attrition = \frac{N_{assigned\_sample} - N_{analytic\_sample}}{N_{assigned\_sample}}$$

The study's student attrition was 27% overall, 26% for treatment students, and 29% for control students. Thus, the student sample was both randomly assigned and had low overall and differential attrition, which provides minimal risk of in-moving students or attrition biasing the treatment estimate (IES, 2017).

Sample size, power, and precision. Power estimates suggested that four students per focal class for each sample would provide minimum detectable effect sizes (MDES) of .15 or below for each analysis. Random sampling was performed from the pretest student population into an assigned sample without regard to whether they completed posttest writing according to a number randomly assigned at baseline. The analytic sample therefore, provides an unbiased sample of those students who remained in the sample between the beginning and end of the study.

Random Assignment Method. Teachers were randomized together with the students in a single focal class, within school-by-grade blocks to provide better equivalence across baseline indicators of key outcomes and of contextual factors. Within cooperating districts, each site recruited approximately 60 grade 7-12 ELA teachers to participate in the experiment. Before randomization, SRI chose a Year 1 focal classroom for each teacher and collected rosters of students enrolled. SRI then randomized approximately half of the recruited teachers (and their classrooms) into treatment and half into control (delayed treatment). Control teachers were asked to maintain business-as-usual conditions and were offered a 1-year version of the program at the end of the study. Randomization occurred within grade-level blocks consisting of school-by-grade groupings, although not all schools had multiple study teachers in the same grade. In

these cases, SRI formed blocks across schools, although teachers were still within the same grade and district.

Description of the Pathway Professional Development Program

As mentioned previously, the PD consisted of 6 full-day and 5 afterschool sessions interspersed throughout the school year. The PD was delivered by each of the Site Directors, all long-standing NWP Directors with professional appointments at universities, assisted by Co-Directors, doctoral students, and/or Writing Project Teacher/Consultants. The delivery of the intervention was staggered so that the site led by the intervention developer could model each PD session which the other three Site Directors could watch, followed by a debrief after each session prior to implementing the PD at their respective sites. (See Scaling Strategies for further information on cultivating reform ownership among the other Site Directors.) Each of the sites served approximately 30 teachers from grades 7-12 in the Year 1 treatment. Training was held at either a hotel or PD center removed from teachers' specific school sites to create a professional atmosphere.

As our theory of action indicates, our PD was designed to train teachers to use a cognitive strategies approach to reading and writing instruction in order to make visible to teachers and, in turn, to students the thinking tools research indicates experienced readers and writers access in the process of meaning construction. The goal was to increase teachers' repertoire of instructional strategies as well as the time they devoted to focusing on higher level interpretation of complex text, on generating text-based analytical writing, and especially on revision which, in turn, would result in students' use

of cognitive strategies in reading and writing, increased student engagement, and improved performance on text-based analytical writing tasks.

Introductory Cognitive Strategies Readers' and Writers' Tool Kit Tutorial. During the first two PD days, teachers were introduced to a model of the cognitive strategies that make up a reader's and writer's mental tool kit in Figure 1. These thinking tools or acts of mind directly map on to the CCSS-ELA standards that call for students to be able to do the following as they read and write about complex texts: summarize, make inferences, analyze, interpret, draw conclusions, evaluate, assess, revise, and reflect. Teachers were given the following analogy as a possible model to introduce the concept of the tool kit:

> When we read, we have thinking-tools or cognitive strategies inside our heads that we access to construct meaning. Researchers say that when we read, we're composing, just as when we write. What they mean is that while we read, we're creating our own draft of the story in our heads and as we keep reading and come across something we didn't expect to happen or suddenly make a big discovery about what something means, we start on a second draft of our understanding. So, when you think of yourself as a reader or writer, think of yourself as a craftsman, skilled in making things with your hands, but instead of reaching into a metal tool kit for a hammer or a screwdriver to construct or build tangible or real objects you can actually see, you're reaching into your mental tool kit to construct meaning from or with words.

[Insert Figure 1 here]

Cognitive Strategies Tutorials. To build students' declarative knowledge (Paris, Lipson, & Wixon, 1983) of what cognitive strategies are, teachers presented scaffolded lessons called "tutorials" (Bruner, 1978) in which they introduced each of the tools in the tool kit within the context of reading and writing about high-interest literary or nonfiction texts. All of the texts we selected for the Year 1 tutorials were related to the idea of paying tribute and focused on topics of respect, resilience, courage, heroism, etc. Our first

tutorial focused on Toni Cade Bambara's "The War of the Wall" (1996), a short story that deals with two African-American teenagers who react strongly to a "painter lady" who comes into their neighborhood and starts "messing" with a wall they consider to be their territory. Unaware of her motives, they instantly form a negative impression of this intruder and devise a plan to retaliate by destroying her creation with graffiti. However, when they arrive one morning, paint cans in hand, to fulfill their mission, they find that the painter has gone and left an awe-inspiring mural celebrating the members of their community and paying tribute to her cousin, who was killed in the Vietnam War.

In order to develop students' procedural knowledge, students were given bookmarks, listing sentence starters for each of the cognitive strategies that illustrate what goes on in the mind of a reader or writer in the act of meaning construction. They used these sentence starters to write marginal annotations on the text and to develop more extended commentaries about key quotes. For example, using the revising meaning sentence starter, a student might write, "At first I thought that the painter lady was an intruder just like the characters did and that she didn't respect the members of the community, but now I see that she was using her artistic talent to celebrate them." This might lead him/her to use the reflecting and relating sentence starter, "So, the big idea I'm getting is not to jump to conclusions about people before you get to know them. I also see that sometimes walls can unite us instead of separate us." (See Figure C in the online supplementary materials for the bookmarks with cognitive strategy sentence starters.)

Following the introductory tutorial, to foster conditional knowledge of when to use a cognitive strategy, which strategy or strategies to use, and why, students were

taught to think aloud in response to complex texts while a partner recorded their responses and then labeled their strategy use, as well as wrote metacognitive reflections describing their process of meaning construction and the strategies they used to form interpretations about texts. This helped them to internalize cognitive strategy use, gave them a common language to talk about their thinking, and helped to transfer control to students as they gained competence and learned to apply strategies independently (Langer & Applebee, 1986).

Formative Assessment and Revision Tutorial. Among the 15 elements of effective adolescent literacy programs, Biancarosa and Snow (2004) have theorized that three are most critical to improving student outcomes: (a) ongoing and sustained professional development to improve teacher practice, (b) the use of pretest student data to inform instructional activities, and (c) the use of summative outcomes to evaluate efficacy. The centerpiece of the intervention is an extensive set of materials sites shared with teachers during two subsequent days of the PD, focused on the revision of the students' pretest writing assessment (a text-based analytical essay) into a multiple-draft essay. Student performance on this timed, on-demand pretest essay was used to inform the intervention as teachers engaged in analyzing students' work and identifying students' strengths and areas for growth (Black & Wiliam, 1998). Two of the 10 CCSS-ELA reading anchor standards focus on determining and analyzing themes in both single and multiple texts. Despite its importance in the CCSS-ELA, many students have an inadequate grasp of what a theme is. For this reason, the writing prompts required students to: 1) analyze two nonfiction texts, newspaper articles "Sometimes, the Earth is Cruel" by Leonard Pitts (2010), a tribute to the resilience of the Haitian people after a

violent earthquake, and "The Man in the Water" by Roger Rosenblatt (1982), a celebration of the selflessness of one man who sacrificed his life to save survivors of a plane crash in the Potomac River; 2) develop a theme statement about the author's message; and 3) evaluate the author's purpose for writing each article.

As part of the PD, teachers received copies of the students' pretests with feedback from trained graduate students and/or Writing Project Fellows and were given time to analyze their students' written work, fill out a graphic organizer to document: what their students did well on their pretests, what they failed to do or struggled with in their essays, and specific steps they, as teachers, might take to prepare them to improve their performance on the posttest. They then met in grade level groups to compare their findings, set goals, and develop a plan of action. This helped to cultivate the reform ownership necessary to implement the revision tutorial. Many students had difficulty comprehending what they were asked to do in the prompt or at least failed to respond adequately to the requirements of the prompt, often failed to articulate a theme statement, relied (sometimes exclusively) on retelling/summarizing what happened in response to the traumatic events in both articles, and demonstrated a lack of sentence variety and weak grasp of sentence boundaries. The revision tutorial focused on the following four pedagogical activities to help students become more strategic essay writers: the DO/WHAT Chart, Theme Video Clips Activity, Essay Color-Coding, and Grammar Brush Strokes. Because the revision tutorial involves comparing and contrasting elements of weak and strong essays to model writing strategies and we did not want to short-circuit the students' own thinking about "Sometimes, the Earth is Cruel" or "The Man in the Water," the revision tutorial focused on another newspaper article, "For One Quake

Survivor, Self-Help in the Face of Seeming Helplessness" by Mark Magnier (2011), a celebration of the heroism of an ordinary citizen, Hideaki Akaiwa, who exhibited an extraordinary act of courage by donning scuba gear and plunging into the aftermath of the Japanese tsunami in order to save his wife from the perilous floodwaters.

DO/WHAT Chart. One of the first challenges that students encountered when beginning to write was making sense of the prompt. How students conceive of and define the "problem" of writing has been shown to have tremendous effect on the writing they produce. To help students navigate the prompts students were taught a planning and goal setting strategy that involved the creation of a DO/WHAT chart, which enabled them to deconstruct a prompt and create a roadmap for composing. To create a DO/WHAT chart, students used green and blue highlighters to mark all of the verbs in the prompt which instructed a student to *do* something in green and underline the task words that tell the student *what* to do in blue. For example, they had to write (green) and essay (blue) and make (green) a claim (blue) about the main point, lesson, or message of the text. They then transferred those words onto a T-chart below the prompt. This activity helped the students to clarify and visualize what is expected, plan and goal set, organize information, and to evaluate the criteria for a successful response to the prompt. (See Figure D in the online supplementary materials for a prompt and DO/WHAT Chart.)

Theme Video Clips Activity. According to Gurney, Gersten, Dimino, and Carnine (1990), even at the high school level, theme is the most difficult of the story elements for students to grasp, and it requires more extensive teacher modeling and direct explanation than the other components. At pretest, a majority of the students in the study neglected to articulate a theme statement or had difficulty distinguishing between a topic

and theme. To help students differentiate between a topic and a theme, teachers identified the topic as the *What* of the story, usually expressed as an abstract noun like *loss, family, courage, resilience, etc.*, and the theme as the *So What* about the topic, consisting of a complete sentence such as "Loss brings people together" or "In dire circumstances, even ordinary people are capable of extraordinary acts of courage." To help students practice creating theme statements, teachers first used clips from familiar movies or music videos to engage students in analyzing and interpreting a more accessible text before attempting to determine a theme in a complex text.

Color-Coding for Revision. Another very common shortcoming in pretest essays was the overreliance on summarizing. As has been widely reported, ELs who have been in ELD programs often receive instruction that focuses primarily on literal comprehension. Consequently, they tend to rely on retelling when writing a text-based analytical essay as a way to prove that they understood what they read rather than offering interpretation and commentary to support their argument. As Scardamalia and Bereiter (1987) point out, novice, inexperienced, and struggling writers use a simplified version of the idea-generation process they call *knowledge-telling*, which consists of retrieving information from long-term memory and converting the writing task into simply regurgitating what is known about a topic. More experienced writers, on the other hand, engage in a complex composing process known as *knowledge-transformation*, in which they analyze the writing task and plan what to say and how to say it in accordance with rhetorical, communicative, and pragmatic constraints. One way to help students move from *knowledge-telling* to *knowledge-transformation* is to help them make their thinking visible after they have composed a first draft of an essay using a color-coding

process. Teachers first designated three colors for the types of assertions that comprise a text-based analytical essay and said the following:

> Plot summary reiterates what is obvious and known in a text. Reiterate means to repeat in order to make something very clear. Plot summary is yellow because it's like the sun. It makes things as plain as day. We need some plot summary to orient our reader to the facts, but we do not need to retell the entire story. Commentary is blue like the ocean because the writer goes beneath the surface of things to look at the deeper meaning to offer opinions, interpretations, insights, and "Ah-Ha's." Supporting detail is green because like the color, it brings together the facts of the text (yellow) with your interpretation of it (blue). It is what glues together plot summary and commentary. It's your evidence to support your claims, including quotations from the text.

The next step was to model the process of color-coding as in the following example:

Yellow	Mark Magnier compares Hideaki Akaiwa to a live action hero.
Blue	He depicts Akaiwa as larger than life and portrays him as a symbol of determination and courage.
Green	As Magnier remarks, "In the face of seeming helplessness, he can't sit still."
Yellow	A man of action, he took control and risked his life to rescue victims of the tsunami.
Blue	In the midst of death and disaster, his act of courage gives us hope and teaches us that if Akaiwa can triumph over adversity so can we.

After students were introduced to the color-coding system, they practiced coding sample essays that were marginal/not pass (1 to 3 on a 6-point scale) and adequate to strong pass (4 to 6 on a 6-point scale) and then applied the color-coding strategy to their own first drafts to visibly see whether they had simply summarized or whether they had provided ample textual evidence and commentary. The coded draft then became a visible guide for revision.

Image Grammar. Crafting sentences for rhetorical effects when writing analytical essays is challenging for all students, but especially for ELs because they are in the process of learning their audience's expectations and developing the linguistic

resources to meet them. One strategy to help students with sentence variety is to teach them image grammar brush strokes (Noden, 2011). During the revision tutorial, teachers were introduced to five specific brush strokes to help students enhance their writer's craft through words, syntax, and punctuation: painting with action verbs, adjectives out of order, appositives, participles, and absolutes. Teachers introduced each brush stroke, modeled it, provided students with multiple opportunities to practice, and then required them to add a brush stroke to their revised pretest essay. An example for appositives is as follows:

Painting with Appositives

An appositive is a noun or noun phrase set off by commas that follows/describes the noun it identifies.

> **Teacher Model:** The tsunami, *a huge and ravenous wave*, attached the shore, swallowing up whole buildings in one gulp.
> **Student's Sample:** The tsunami, *a giant monster*, was "picking up cars like they were toys and destroying buildings like they were paper."

To reinforce the tutorial, we developed a website that supported students in composing sentences using distinct grammar brush strokes. After composing, students were matched with peer-reviewers who provided anonymous feedback on their sentences. Students then submitted revised versions of their sentences to their teachers via the website. Since its inception, approximately 545 teachers and 7,136 students have registered with the site, and students have composed approximately 47,079 sentences and submitted them for peer review using the system.

Analyzing Student Work and Reflecting upon the Impact of the Intervention.

In late April or early May, teachers administered the posttest. Prior to attending the final PD and turning the assessments in for scoring, teachers gave students a copy of their

pretests, asked them to compare and contrast their pre and post essays, and gave them a reflection sheet on which to record what they noticed about their growth over time as readers and writers. This helped to develop students' self-efficacy and engender a sense of responsibility for their learning. Teachers then reviewed students' reflection sheets and wrote their own assessment of the impact of the intervention upon their students' growth over the school year. Meeting in cross grade-level school teams with their Site Team Lead, each school then set goals for Year 2 of the project and developed an action plan.

Description of Business-as-usual Professional Development Activities

Control teachers conducted business as usual using the district English language arts textbook and core novels for teaching. Since Houghton Mifflin Harcourt donated class sets of the *Collections* textbook series as part of the i3 matching program, both treatment and control teachers attended a half day PD session devoted to introducing the *Collections* textbook. Several districts also conducted PD on district benchmark assessments and the new Smarter Balanced Assessment Consortium (SBAC) standardized tests.

Scaling Strategies to Expand and to Replicate the Intervention

Cultivating Reform Ownership

Since the largest intervention studies conducted by the study site prior to the i3 Validation project involved a maximum of 100 teachers in one school district randomized into treatment and control conditions and included approximately 1,750 treatment and 1,750 control students in any given year, the expansion of the project to four NWP sites training 230 teachers in seven districts and serving approximately 4,000 treatment and 4,000 control students considerably increased the number of users or *spread* of the

project. However, Coburn (2003) asserts that "expanding a reform to multiple settings is a necessary but insufficient condition for scale," adding that "scaling up not only requires spread to additional sites but also consequential change in classrooms, endurance over time, and a shift such that knowledge and authority for the reform is transferred from external organizations to teachers, schools, and districts" (p. 4). We would argue that a key missing ingredient in Coburn's dimensions of scale is the necessity of establishing reform ownership among those deliverers of the intervention who were not the original developers of the program. Importantly, the study site recruited three other NWP Site Directors in southern California to participate in the intervention, enabling us to build upon our long-standing pre-existing relationships, substantial shared knowledge about writing instruction, and common stance towards teacher PD. Prior to the launch of the i3 PD the Site Directors and their implementation teams consisting of key Teacher Leaders from their sites met for two full days, experienced the cognitive strategies intervention PD as participants, and discussed each component of the intervention at length. Another advantage in our partnership was that we were geographically within driving distance of one another. As mentioned previously, this allowed us to develop a staggered PD delivery schedule in which the study site PD for its school district was held a week or two in advance of the other sites. The other Site Directors and their teams attended, took notes, and debriefed after each full day session, asking questions, sharing ideas, and making suggestions. Klein, McArthur, and Stecher (1995) discuss the concept of "alpha" sites that devise and refine innovations and then replicate them in multiple sites. In our collaboration, the study site was referred to, somewhat in jest, as the "mother ship," and, from the outset, the understanding was that the other Site Directors, especially given their

expertise, were not expected to be "sock puppets." This stance did create the possibility of a tension between fidelity and reform ownership (Coburn, 2003). So, certain "non-negotiables" were established. Each Site Director agreed to: 1) teach the introductory cognitive strategies tutorial, making minor modifications, if needed; 2) train readers to respond to treatment students' pretests and provide those papers with feedback to the treatment teachers and their students; 3) teach the revision tutorial designed to guide treatment students through revising their pretests, also with modifications, if needed; 4) score the treatment and control students' pre/posttests on a study site prompt-specific rubric (independent from the rubric used to measure treatment impacts); and 5) return those pre/post scores to teachers at the beginning of Year 2. However, Blumenfeld and colleagues (2000) have noted that "for instructional reforms to be sustainable across settings, they must be tailored to those settings. As a result, enactments may vary and still be warranted with respect to the basic principles of the innovation" (p. 162). Therefore, while all three other Site Directors delivered the essentials of the treatment with fidelity, their differing leadership styles, the specific district settings and teacher profiles, and the degree to which those teachers had prior exposure to PD all influenced how the treatment was implemented. Based on the interviews with Site Directors, observations of the PD and teacher practice, SRI concluded that one site sought to replicate the intervention as faithfully as possible, consciously limiting any adaptations in Year 1. Another site included two classroom teachers not-employed in the treatment district on their implementation team. After each PD session, these teachers taught the most recent PD strategies in their classrooms and used their own experiences to suggest modifications or additions in advance of their delivery of the PD to their school district. The third site

adapted the model by incorporating teacher presentations on topics of their choice throughout the PD.

As McLaughlin and Mitra (2001) point out, a shift in reform ownership requires transferring substantive and strategic decision making from the reform organization to district and school leadership. Hence, throughout Year 1, we endeavored to engender in district personnel and teachers a sense of ownership in and responsibility for spreading and sustaining the intervention over time. In addition to meeting with district partners as well as school principals to learn about and respond to or align with district and school site initiatives, each Site Director also met with the ELA teachers prior to the start of the project. To enlist teacher buy-in, teachers were surveyed regarding their interests and most pressing needs in regards to ELA instruction and serving the ELs in their classrooms. Their responses were used to construct the year-long plan for the afterschool workshops. Hence, each NWP site had the flexibility to customize the afterschool portion of the program.

Striving for Depth and Sustainability in Teachers' Instructional Shifts

Coburn (2003) notes that to be "at scale," reforms must "effect deep and consequential change in classroom practice" (p. 4) and these changes must be informed by pedagogical principles that are embodied in the enacted curriculum. Strategy instruction in this intervention occurs within the context of teaching reading and writing as a process and involves prereading, during reading, and postreading activities, as well as prewriting, planning, drafting, sharing, revising, and editing activities. Teachers were encouraged to slow down and go deep and to engage students in reflecting upon their cognitive strategy use and their growth as readers and writers over time. This

instructional approach necessitates more time and attention devoted to reading and writing and, consequently, generates more workload for both teachers and students. The more ambitious a reform is and the more effort it demands, the more it will need "well-elaborated materials and sustained ongoing professional development to achieve depth" (Coburn, 2003, p. 9; see also Sternberg et al., 2011). The intervention provided a wide array of well-elaborated, teacher-tested and easy-to-use paper and computer-based instructional materials. Because these materials were designed to be presented to students across the grade levels (7 to 12) and with varying degrees of language proficiency, teachers were given time to meet in grade-level groups and as school teams to discuss how to modify the materials to meet their specific students' needs. To increase the likelihood that the intervention would be sustained over time, each participating school designated one exemplary teacher as the Site Team Lead who coordinated efforts at the school site, met with NWP Site Directors to provide feedback on how teachers and students were responding to the intervention, to provide coaching in Year 3 when the control teachers received the treatment, and eventually to become school site literacy coaches after the termination of the grant funding. Coburn (2003) points out that to achieve true depth and sustainability, "it is important to look beyond the presence or absence of specific materials or tasks to the underlying principles embodied in the way teachers engage students in using these materials and tasks" (p. 5). Intervention strategies and tutorials were presented not as a curriculum but as models of an instructional approach to providing thinking tools that empower students to become active participants in their own learning. Teachers were invited to modify, supplement, and add their own innovations to these materials as they worked at their school sites with Site Team Leads.

Hence, reform ownership, depth, and sustainability exist in a synergistic relationship with one another.

Data Sources and Collection

SRI collected data from multiple sources to understand program implementation across the four Writing Project sites, measured differences in PD experienced by treatment and comparison teachers and in teacher practice through a teacher survey, and assessed student learning with a study-administered on-demand student writing assessment and state standardized test scores in ELA.

Program Implementation. To assess the fidelity of program implementation across the sites, SRI worked with the intervention developers to develop indicators for four key program components: Duration and breadth of ***teacher participation*** in PD; ***Content*** of the teacher PD***; Scaffolding*** approaches for teachers***;*** and ***Formative feedback*** mechanisms.

SRI collected information on teachers' participation in PD and the features of the PD and examined results against predetermined thresholds for each indicator. SRI aggregated from the teacher level to the Writing Project site level to describe the proportion of teachers in each site attending a sufficient number of PD events to meet the agreed-on thresholds (e.g., 90% of teachers participated in at least four of five full-day PD events). SRI also reviewed other data sources, most notably artifacts from and observations of the PD to determine whether the sites provided the content and strategies for instructional delivery consistent with the intervention model. In examining artifacts, SRI noted the presence or absence of specific indicators. For example, sites were expected to present teachers with two intervention "tutorials"—an introductory tutorial on cognitive

strategies and a revision tutorial designed to teach students a process for revising their writing. SRI analysts reviewed PD agendas, PowerPoints, and handouts for evidence that sites met this expectation. Through observing key PD events at each of the intervention sites, researchers were able to independently verify the information recorded on the artifacts as well as to note fidelity and adaptations to the original model.

SRI also sought to understand the contrast between treatment and control teachers' PD experiences via an annual teacher survey administered to all treatment and control teachers (with an 88% response rate). The items asked about the hours spent in PD focused on writing instruction and about the nature of the ELA-focused PD teachers participated in. Specific items asked about whether the PD supported Common Core implementation, focused on ELs, and reflected the key features of the intervention PD. The survey also included questions on the extent to which the ELA-focused PD emphasized intervention-aligned approaches to reading and writing instruction. Items describing this contrast were analyzed using the same methodology as that estimating impacts on teacher practice (see below).

Teacher Practice Outcomes. SRI measured teacher practice through a survey of treatment and control teachers in spring, using a construct modeling approach to identify and develop the survey, which was mapped to the intervention's theory of action. The survey included 50 items about teachers' allocation of instructional time, frequency of teaching specific cognitive strategies when teaching reading, and frequency of teaching specific aspects of the writing process when teaching writing (e.g., composing a thesis statement, sentence craft and variety, balancing the use of summary, supporting detail, and commentary). SRI also interviewed 33 teachers across the four sites about the impact

of the intervention on their teacher practices using a semi-structured protocol. Further, each NWP site conducted interviews with treatment teacher focus groups at the school site level, also using a semi-structured protocol, to determine what aspects of the PD were most effective in impacting teacher practices. *Impact Estimates.* Models estimating teacher outcomes were multi-level models similar to those estimating student impacts (described below), although the student level and baseline covariate were omitted. The models also included grade and district-fixed effects.

Replication and Scaling Processes and Structures. To explore what processes and structures supported replication and scaling up of the Pathway intervention, SRI conducted interviews with the four Site Directors and 7 principals across the sites to investigate: the supports that facilitated replication and the challenges that new sites faced. In addition, the lead site collected written surveys from all NWP partners regarding how each site developed reform ownership among the district and teachers, adapted the intervention to meet specific district needs, addressed and overcame challenges, and created plans for sustainability. These were synthesized into a year-end report submitted to i3.

Student Learning Outcomes. SRI examined two types of student learning outcomes: one was writing data from study created prompts and the other were standardized assessments. *Writing Data.* SRI administered a 2-day writing assessment to all students in focal classrooms in the fall as a baseline measure and again in spring as an outcome. On day 1 the prompts provided students with one literary nonfiction text to read and several activities designed to scaffold their analysis of the text. On day 2 students were asked to hand-write an analysis of the text. Prompts were counterbalanced such that

teachers were assigned one of two prompts in the fall and the other in the spring. Teachers administered the baseline prompt before learning of their randomization status and before the first PD session. Both the analytical reading and the text-based writing are aligned with Common Core Standards, and the performance tasks are similar to those included on some state assessments (e.g., Connecticut) and the national consortia assessments (i.e., PARCC and Smarter Balanced).

The writing assessment was used for both formative and summative purposes. Treatment students revised their baseline writing and the teachers received pre/posttest scores using the study-and prompt-aligned rubric as part of the Pathway PD. SRI also provided de-identified prompts to the National Writing Project (NWP) to score on a prompt-agnostic rubric for evaluation purposes. The student writing was scored with the Analytic Writing Continuum for Literary Analysis (AWC-LA). Over a decade, the NWP developed the Analytic Writing Continuum (AWC), which has been shown to be a valid and reliable measure of student writing (Bang, 2013). The NWP worked with a panel of writing assessment experts to modify the AWC to more accurately score literary analysis, with a focus on the development of ELs' writing. Each handwritten paper was given a holistic rating as well as ratings on each of four attributes: content, structure, sentence fluency, and conventions. Scorers did not know the students' treatment status or the time at which the paper was collected (pre- or posttest). Scorers (N=112) were recruited from current and former teachers affiliated with local NWP sites not participating in the intervention. SRI monitored the scoring to ensure the NWP followed impartial processes. Reliability of the prompt scoring was assessed separately for each writing attribute measure in the AWC-LA through double scoring 10% of the papers. Raters agreed within

a single score point for 90% of papers on the holistic score and structure attribute, 91% on the content attribute, 88% on the sentence fluency attribute, and 87% on the conventions attribute.

Data Analytic Strategy

Primary Student Impact Estimates. To measure the impact of the Pathway Project on student writing as measured by the NWP AWC-LA, SRI undertook two sets of calculations. The first compared the pretest to posttest gains of Pathway and control groups for the total analysis sample and for demographic subgroups. The second used the following gain score regression analysis model. The predicted writing ability for student *i*, in teacher *j*'s classroom, in blocking pair *k* as a function of being assigned to treatment is given as:

$$Y_{ijk} = \beta_0 + \beta_i(Treatment_k) + \beta_2(\overline{Pre-test_i}) + \boldsymbol{\sigma_k} + \epsilon_{ijk} + \mu_{jk} + \eta_k$$

Impacts on the AWC-LA were estimated separately for the holistic score and each of the individual analytic scores. District-fixed effects $\boldsymbol{\sigma_k}$ account for variation in policy or achievement by district. Random effects ϵ_{ijk}, μ_{jk}, and η_k allow for error at the student, teacher, and randomization block level, respectively. Randomization block is entered as a model to account for the structure of random assignment, wherein teachers were randomized within blocks, as described above. The model includes a control for student-level baseline achievement on the same baseline score (e.g., holistic AWC-LA baseline scores when predicting the holistic AWC-LA outcome). Student baseline and outcome scores are standardized within cohort and prompt form to account for prior achievement, cohort at baseline, and prompt effects. β_1 provides an estimate of the effect of teacher assignment to treatment on student writing performance (the Intent-to-Treat effect),

standardized within the analytic sample. SRI estimated these multilevel models using the Stata 14.2 *mixed* command. They used restricted maximum likelihood estimation and the Kenward-Roger method to compute degrees of freedom for the models and calculate *p*-values to adjust for sample sizes at the teacher and block level (Kenward & Roger, 1997; Schaalje, McBride, & Fellingham, 2002). In addition to providing student-level descriptive statistics of pretest data for the analytic sample, SRI formally established baseline equivalence for the analytic sample by predicting the pretest data using the same structural model used to predict the outcome scores to estimate the treatment effect. This model predicted standardized baseline differences on the holistic AWC-LA of -.02 ($p>.1$), meeting the What Works Clearinghouse standards for baseline equivalence (IES, 2017).

State Assessment Data. As an additional measure of student achievement, SRI collected state ELA assessment data as available. California began administering the Smarter Balanced Assessment Consortium (SBAC) assessment in spring 2014 to students in grades 3–8 and 11. These data served as an additional outcome measure for students. As California did not require a standardized ELA assessment during the 2013–14 school year, SRI used the spring 2013 California Standards Assessment (CST) as a baseline measure for those students randomized at the study's onset. Given the availability of both outcome and baseline data, these impacts are estimated for all students in grades 7–8 and 11.

Additional Student Impact Estimates. In addition to the estimates described above, SRI provided several additional impact estimates, including the moderating effects of the programs by subgroups of students and impacts on students using the SBAC assessment

as an outcome. Impact models are estimated similarly to those above. To understand whether the intervention had differential impacts on subgroups of students on the AWC-LA, SRI ran three analyses: impacts for ELs compared with non-ELs,[1] for Latino students compared with White students, and for girls compared with boys. These models are similar to the student outcome models described above, although the analytic sample size changed (dropping to students with non-missing data categorized into one of the two groups). In addition to the terms described in the model above, an indicator variable was included for the subgroup and an interaction term was included between the subgroup and treatment indicator. Finally, SRI ran impacts on California's standardized ELA assessment, SBAC. SBAC and CST data were standardized within the analytic sample for each grade.

Results

Here we describe our findings relating to program implementation teacher practice, scaling up the intervention, and student learning outcomes.

Program Fidelity

The intervention program was implemented largely as intended, though teacher participation fell short of pre-determined thresholds.

- For the ***teacher participation*** component, fidelity of implementation was defined as 90% of teachers participating in (a) at least four of the five full-day PD events and (b) at least three of five after-school PD events. Three sites met the threshold for full-day events and two met the threshold for after-school events.
- Fidelity of implementation for the ***content*** component was defined as presenting teachers with two intervention "tutorials" that they were to implement in their classrooms—an introductory tutorial on cognitive strategies tool kit and a revision tutorial designed to teach students a process for revising their writing—and

[1] English learners included both students currently receiving English learner services and those who had previously received services but had subsequently been re-designated as Fluent English Proficient. The reference group included both students who only spoke English and those with a second language who had never received English learner services.

setting the expectation that teachers would take their students through the process of revising their own work (i.e., a multiple draft essay). To ensure that all treatment teachers' students completed the pretest revision, teachers were required to submit the revised essays along with the posttests in order to qualify for their stipends. All four sites met implementation thresholds for content.

- The ***scaffolding*** component involved modeling lessons for teachers, in particular the introductory and revision tutorials, and providing teachers with all of the instructional materials they would need to go back to their classrooms and teach the intervention lessons (including, for example, pre-printed class sets of readings and handouts). Again, all sites met implementation thresholds for scaffolding.
- The ***formative feedback*** component involved the Writing Project sites working with pre-service teachers at their universities to read and provide feedback on students' "pretest" writing. To meet this threshold, all sites were to provide each teacher with full class-sets of commented-upon student papers. Three of the four sites met this threshold. The formative assessment component also involved providing all teachers with data on their students' gains from the pre- to posttest in Year 1 via "teacher results letters" that teachers received at the beginning of Year 2. All sites met this threshold.

SRI also examined the contrast between treatment and control teachers' experiences with ELA-focused PD and found large, statistically significant differences between the treatment and control group in the amount, content, and scaffolding of the PD received. Treatment teachers reported receiving more ELA-focused PD (41 hours) than control teachers (9 hours), and a greater emphasis on cognitive strategies in reading, as well as on specific writing strategies and skills, such as balancing the use of summary, supporting detail, and commentary, composing a thesis statement or developing a claim, and sentence craft and variety. Treatment teachers were more likely than control teachers to report that the ELA-focused PD they participated in included the scaffolding that is central to the intervention model, such as providing teachers with prepared lessons and class-sets of materials designed to support classroom implementation and analyzing student work to inform instruction. (See Table B in the online supplementary materials.)

Teacher Practice Outcomes

In Year 1 85% of treatment teachers reported implementing the cognitive strategies tutorial and 92% reported using the revision tutorial; however, treatment and control teachers reported spending equal amounts of time teaching analytical essay writing and reading strategies. Middle school treatment teachers and lower grade high school teachers (grades 9-10) tended to implement the intervention with greater fidelity than upper grade high school teachers (11-12). (See Tables C and D in the online supplementary materials.)

Scaling Up Outcomes

Based on the interviews with each of the Site Directors at the three replication sites, the following processes and structures supported the replication and scaling up of the intervention: 1) the pre-existing relationship between the four NWP Site Directors; 2) the PD provided to the three replication Site Directors before the launch of the PD; 3) the modeling provided by the lead site that Site Directors watched and debriefed on before delivering the PD themselves; 4) the ready-to-use materials developed and disseminated by the lead site for replication sites to implement. In terms of Coburn's model of scaling up, SRI found that: the site which sought to replicate the model as faithfully as possibly, consciously limiting adaptations, as a result, may have also limited their overall reform ownership; the site which actively sought to make the intervention their own by having teachers try out the activities in their classrooms and suggest modifications, prior to the site delivering the PD, demonstrated the greatest reform ownership. Finally, the site that adapted the intervention by incorporating teacher presentations of their choice throughout the project did not necessarily increase reform ownership when their adaptations were not in the furtherance of the Pathway model. As with teacher implementation, the premise

behind a multiyear PD design is that reform ownership builds over time and that both replication Site Directors and teachers need more than one year to internalize the intervention and make it their own.

Student Learning Outcomes

Writing Assessment

The intervention increased the performance of students in the treatment group nearly a third of a standard deviation relative to the performance of students in the control group (Table 2, $p < .001$). Results were positive and statistically significant for not only the holistic score (d=.32), but all four of the analytic scores: content (d=.31), structure (d=.29), fluency (d=.27), and conventions (d=.32).

[Insert Table 2 here]

SRI found no evidence that the intervention effects were moderated by English learner status, gender, or ethnicity (Table 3). Impacts were positive and statistically significant for all groups in each model. Tests of difference between the two effect sizes found no evidence of difference between the impacts for any of the subgroups. However, in examining posttest results, we note that Latino treatment students achieved parity with their White peers in the control condition (Tx Hispanics, 3.3; Ctl White, 3.3). Further, treatment ELs and RFEPs achieved parity with their control non-EL and IFEP peers (Tx ELs/RFEPs, 3.2; Ctl non-ELs/IFEPs, 3.2).

[Insert Table 3 here]

State Standardized Test Scores

SRI found no evidence of a transfer of the intervention's impact on writing to a broader measure of student ELA achievement (*d*=.06, s.e.=.05, p>.10, *dof*=83, student *n*=2,733, teacher *n*=134, block *n*=51).

Discussion

As the report "Scaling up Evidence-Based Practices: Strategies from Investing in Innovation (i3)" (DeWire, McKithen, & Carey, 2017) points out, "broadening the reach of evidence-based innovations, or scaling up has been a challenge in education improvement efforts" (p. 1). In their review of 67 programs funded by i3 which were completed by May 2017, Abt Associates (Boulay et al., 2018) found that only 12 of the 67 impact evaluations (18%) yielded a statistically significant positive impact on at least one student achievement outcome. However, Validation and Scale Up grants (i.e., those funded based on prior evidence of efficacy) were more likely to improve a student academic outcome than interventions supported by Development grants (which required no such evidence). Of the 31 evaluations of student achievement in English language arts, 7 yielded statistically significant results, all of which were positive. The effect sizes for other ELA interventions for secondary students were: Collaborative Strategic Reading (*d*=.04); Reading Apprenticeship (*d*=.14); College Ready Writers Program (CRWP) (*d*=.20); and the Expository Reading and Writing Course (ERWC) (*d*=.13). Hence, our main finding that scaling up the intervention to four National Writing Project sites in southern California yielded positive, significant treatment effects on the holistic measure (*d*=.32) and on all four analytic categories of content, structure, fluency, and conventions (average *d*=.29) compares favorably with these i3 studies and provides comparable

results with the Year 1 results of the UCI Writing Project's original randomized controlled field trial in SAUSD (*d*=.35) (Kim et al., 2011).

Of the i3 ELA evaluations reported to date, none were targeted for English learners. Given the dearth of research regarding effective literacy interventions for predominately Latino secondary mainstreamed ELs, we believe our work has the potential to contribute to the scientific knowledge base regarding strategies for enhancing the academic literacy of secondary ELs both in California, which serves 29% of the nation's ELs enrolled in K-12 schools, and nationally (Sanchez, 2017). Our findings highlight the efficacy of implementing a cognitive strategies approach for ELs using a range of pedagogical strategies to make visible for ELs the thinking tools accessed by experienced readers and writers during the process of meaning construction. Further, they demonstrate that such an approach can be successfully delivered by program facilitators not previously involved in the intervention development.

As opposed to the UCI Writing Project's OELA-funded randomized field trial in Anaheim Union High School District, where the intervention conferred the greatest benefits to ELs, RFEPs, and Latino/Hispanic students (Olson et al., 2017), results from the i3 scale up study show no evidence that the intervention had differential effects by English learner status, gender, or ethnicity. However, results for all of the sub-groups were positive and statistically significant, and, as was mentioned previously, Latino treatment students achieved parity with their White peers in the control condition and treatment ELs and RFEPs achieved parity with their control non-ELs and IFEP peers. Hence, the intervention contributed to reducing the achievement gap in writing for these groups. SRI's analysis also revealed no impact upon the SBAC standardized test as

opposed to a small but statistically significant impact which was achieved in our IES RCT in SAUSD on the California Standards Test (CST) (d=0.10) (Kim et al., 2011; Olson et al., 2012). Since the SBAC was administered for the first time in Spring 2015 on a voluntary basis, this null finding is perhaps not unexpected.

Additionally, our study results are consistent with Englert et al.'s (1991) findings on the influence of teachers who encourage and develop students' cognitive engagement at the elementary level and confirm De la Paz and Graham's (2002) findings about the positive impact of strategy instruction on student writing at the secondary level. In particular, our results indicate that teachers can learn to engage Latinos and mainstreamed English learners in higher level interpretive reading and analytical writing about texts through direct strategy instruction, modeling of strategy use, and creating opportunities for students to practice and apply these skills through teacher coaching and feedback, all practices recommended by an expert panel in the IES Practice Guide *Teaching Secondary Students to Write Effectively* (Graham et al., 2016).

Limitations and Directions for Future Research

Our study does have limitations related both to the reasonably close physical proximity of the scale up sites to the original intervention site and the philosophical proximity of "ideas and beliefs" (Stenberg et al., 2011) between the "creators of the program to its first adopters and implementers" (p. 9). Further research is needed to determine whether the intervention can be successfully scaled up to a national level in more diverse geographical areas with a range of student populations and language backgrounds, and by implementers with less exposure to the National Writing Project. Another limitation is that the fidelity of teacher implementation was assessed via a self-

report survey which was supplemented by 33 semi-structured teacher interviews. To more thoroughly document teachers' uptake of the intervention and the impact on their practices, classroom observations of both treatment and control teachers at several time points would provide a more reliable source of data from which to draw conclusions.

Conclusion

Given that in a recent meta-analysis of effect sizes in relation to methodological features, Cheung and Slavin (2016) found that effect sizes of evaluations of interventions targeting student outcomes in reading, math, and science tended to be smallest in randomized control trials and studies with larger sample sizes, we believe that achieving an effect size of .32 on a large scale writing intervention has important implications for practitioners, intervention developers, and policymakers. Further, our results provide rigorous experimental evidence of the efficacy of the intervention for improving student outcomes for all learners, as well as helping to level the playing field for Latinos and mainstreamed ELs, groups facing persistent educational challenges.

Our study also contributes to the growing body of knowledge regarding how to go beyond spread to genuinely cultivate reform ownership, depth, and sustainability and highlights the need for high quality, sustained PD in order to achieve this goal. Coburn (2013) concludes that because the "the problem of scale remains one of the most pressing issues in educational reform and improvement" (p. 8), conditions are "ripe" for studies that further explore the "multidimensionality of scale" (p. 9). Noting that "broadening the reach of evidence-based innovations—or scaling up—has been a challenge in education improvement efforts" (DeWire, McKithen, & Carey, 2017. p. 1), the Investing in Innovation (i3) program set as one key goal supporting organizations in expanding

research efforts that yield useful information about scaling up evidence-based interventions across diverse educational settings and disseminating lessons learned. Our Validation study was dedicated to meeting this goal. Since in Year 1 of our Pathway intervention, both treatment and control teachers reported spending roughly equivalent amounts of time teaching writing and reading strategies, perhaps reflecting the heavy emphasis on text-based instruction in the CCSS-ELA, we hypothesize that the following practices contributed to the successful replication and scale up of the Pathway intervention, thereby improving the quality of teachers' instruction in order to achieve the statistically significant student outcomes reported in this study:

- building capacity among the NWP Site Directors to deliver the PD prior to the launching of the intervention;
- designing and implementing PD of an intensity and duration to achieve a depth of understanding that would effect change in teachers' practices over time;
- staggering the PD sessions so that the lead NWP site could model PD delivery for the replication sites;
- presenting certain "non negotiables" to NWP replication sites and to treatment teachers in terms of the use of certain strategies, materials, and tutorials to ensure fidelity;
- cultivating reform ownership by inviting participants to modify and adapt project materials to suit specific district, teacher, or student needs, and to share their adaptations with others;

- analyzing student work for both formative and summative purposes and engaging teachers and their students in reflecting upon student growth as readers and writers over time;
- planning for sustainability by grooming Site Team Leads at each school to serve as liaisons and school site coordinators during the treatment and literacy coaches after the termination of the grant;
- viewing the intervention itself as curricular approach, rather than as a set curriculum, that seeks to continuously improve based on input from NWP site, district, and school (teacher and student) partners.

Acknowledgments

We gratefully acknowledge funding from the U.S. Department of Education's Investing in Innovation (i3) program, under grant number U411B130029. The views expressed in this paper reflect the opinions of the authors and not the funding agency or the authors' respective institutions.

Table 1: NAEP Reading and Writing Scores

Grade 8	Writing	Reading	Grade12	Writing	Reading
White	34%	44%	White	35%	46%
Hispanic	14%	21%	Hispanic	11%	25%
Black	11%	16%	Black	9%	17%

Table 2: Impacts on AWC-LA, in Effect Sizes

	Point-Estimate		se	dof	Block N	Teacher N	Student N
Holistic	0.32	***	0.06	163	81	225	894
Content	0.31	***	0.07	163	81	225	894
Structure	0.29	***	0.07	163	81	225	894
Fluency	0.27	***	0.07	163	81	225	894
Conventions	0.32	***	0.07	163	81	225	894

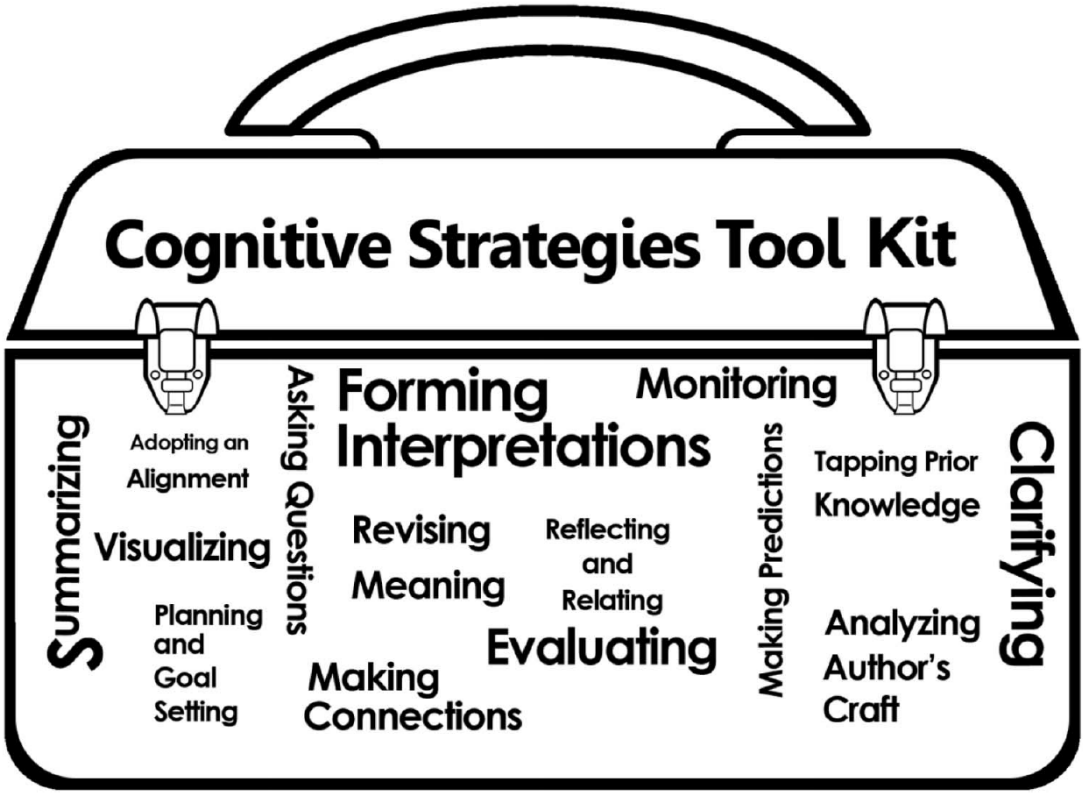
~$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$; Confirmatory contrasts; significance levels represented by asterisks are unchanged by Benjamini-Hochberg multiple comparison correction

Table 3: Impacts on AWC-LA Holistic Scores by Subgroup, in Effect Sizes

Gender		
Effects for boys	0.31	***
(*SE*)	(0.09)	
Effects for girls	0.30	**
(*SE*)	(0.09)	
Difference in Effects Between Boys & Girls	-0.02	
(*SE*)	(0.12)	
Student *n*/ Teacher *n*	856/222	
English Learners		
Effects for English Learners (EL and RFEP)	0.38	***
(*SE*)	(0.12)	
Effects for Non-English Learners (EO and FEP)	0.30	***
(*SE*)	(0.08)	
Differences in Effects Between EL and Non-EL Students	0.08	
(*SE*)	(0.12)	
Student *n*/ Teacher *n*	858/223	
Ethnicity		
Effects for White students	0.43	**
(*SE*)	(0.14)	
Effects for Latino/a students	0.34	***
(*SE*)	(0.08)	
Differences in Effects Between White & Latino/a students	-0.09	
(*SE*)	(0.16)	
Student *n*/ Teacher *n*	740/217	

Note: All samples exclude students for whom subgroup data are missing or who do not fall into one of the named categories. EL includes English learner and redesignated fluent English proficient students. Non-EL includes English-only students and fluent English proficient students. ~p < .10; *p < .05; **p < .01; ***p < .001.

Figure 1: Cognitive Strategies Tool Kit



References

Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college.* Washington, DC: Office of Educational Research an Improvement, U.S. Department of Education.

August, D. E., & Shanahan, T. E. (2006). *Developing literacy in second language learners: Report of the National Literacy Panel on Language-Minority Children and Youth*. Mahwah, NJ: Erlbaum.

Bambara, T.C. (1996). *The war of the wall.* New York, NY: Random House.

Bang, H. J. (2013). Reliability of National Writing Project's Analytic Writing Continuum Assessment System. *The Journal of Writing Assessment, 6*(1). Retrieved from http://journalofwritingassessment.org/article.php?article=67

Batalova, J., Fix, M., & Murray, J. (2005). *English language learner adolescents: Demographics and literacy achievements. Report to the Center for Applied Linguistics.* Washington, DC: Migration Policy Institute.

Batalova, J., & McHugh, M. (2010). *Top languages spoken by English language learners nationally and by state.* Washington, DC: Migration Policy Institute.

Biancarosa, G., & Snow, C.E. (2004). *Reading next—A vision for action and research in middle and high school literacy: A report to Carnegie Corporation of New York* (2nd ed.). Washington, D.C.: Alliance for Excellent Education.

Black, P., & William, D. (1998). *Inside the black box: Raising standards through Classroom assessment.* Department of Education and Professional Studies: Granada Learning.

Blumenfeld, P., Fishman, B., Krajcik, J.S., Marx, R.W., & Soloway, E. (2000). Creating

usable innovations in systemic reform: Scaling-up technology-embedded project-based science in urban schools. *Educational Psychologist, 35*, 149 – 164.

Block, C. C., & Pressley, M. (Eds.). (2002). *Comprehension instruction: Research-based practices*. New York, NY: Guilford Press.

Boulay, B., Goodson, B., Olsen, R., McCormick, R., Darrow, C., Frye, M., Gan, K., Harvill, H., & Sarna, M. (2018). *The Investing in Innovation Fund: Summary of 67 Evaluations: Final Report* (NCEE 2018-4013). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Bruner, J. (1978). The role of dialogue in language acquisition. In A. Sinclair, R. J. Jarvella, & W. J. M. Levelt (Eds.), *The child's conception of language* (pp. 241-256). New York: Springer.

Bunch, G.C., Kibler, A., & Pimentel, S. (2012, January). *Realizing opportunities for English learners in the Common Core English Language Arts and Disciplinary Literacy Standards.* Paper presented at the Understanding Language Conference, Stanford University, CA.

Capps, R., Fix, M., Murray, J., Ost, J., Passel, J., & Herwantoro, S. (2005). *The new demography of America's schools: Immigration and the No Child Left Behind Act.* Washington, DC: Urban Institute.

Cheung, A.C.K., & Slavin, R.E. (2016). How methodological features affect effect size in education. *Education Researcher, 45*(5), 283-292.

Coburn, C.E. (2003). Rethinking scale: moving beyond numbers to deep and lasting change. *Educational Research, 32*(6), 3-12.

Council of Writing Program Administrators, National Council of Teachers of English, & National Writing Project. (2011). *Framework for success in postsecondary writing.* Retrieved from wpacouncil.org/framework

De La Paz, S., & Graham, S. (2002). Explicitly teaching strategies, skills, and knowledge: Writing instruction in middle school classrooms. *Journal of Educational Psychology, 94*, 687-698.

DeWire, McKithen, C., & Carey, R. (2017). *Scaling up evidence-based practices: Strategies from Investing in Innovation (i3)*. Washington, D.C.: Investing in Innovation Community.

Echevarria, J., Short, D., & Vogt, M. E. (2012). *Making content comprehensible for English language learners: The SIOP Model* (4th ed.). Boston, MA: Allyn & Bacon.

Englert, C. S., Mariage, T. V., & Dunsmore, L. (2006). Tenets of sociocultural theory in writing instruction research. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 208–221). New York, NY: Guilford Press.

Englert, C.S., Raphael, T.E., Anderson, L.M., Anthony, H.M., & Stevens, D.D. (1991). Making writing strategies and self-talk visible: Cognitive strategy instruction in writing in regular and special education classrooms. *American Educational Research Journal, 28*, 337-372.

Fitzgerald, J. (2017). How practice-based research informs adolescent English language

learners-composing and compositions. In K. A. Hinchman & D. Appleman (Eds.), *Adolescent literacies: A handbook of practice-based research* (pp. 357-378). New York: Guilford.

Fitzgerald, J., & Amendum, S. (2007). What is sound writing instruction for multilingual learners? In S. Graham, C.A. McArthur, & J. Fitzgerald (Eds.), *Best practices in writing instruction* (pp. 289-307). New York: Guilford.

Flower, L., & Hayes, J.R. (1980). The dynamics of composing: Making plans and juggling constraints. In L. Gregg & E. Steinberg (Eds.), *Cognitive processes in writing* (pp. 31-50). Mahwah, NJ: Erlbaum.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication, 32,* 365–387. http://dx.doi.org/10.2307/356600

Fredericksen, C.H., & Dominic, J.F. (1981). Introduction: Perspectives on the activity of writing. In C.H. Fredericksen, & J.F. Dominic (Eds.), *Writing: The nature, development and teaching of written communication* (Vol. 2, pp. 1-20). Hillsdale, NJ: Erlbaum.

Graham, S. (2018). A writer(s)-within-community model of writing. In C. Bazerman, A. Applebee, V.W. Berninger, D. Brandt, S. Graham, J.V. Jeffery, P.K. Matsuda, S. Murphy, D.W. Rowe, M. Schleppegrell, & K.C. Wilcox (Eds.), *The lifespan development of writing* (pp. 272-325). Urbana, IL: NCTE.

Graham, S., Bruch, J., Fitzgerald, J., Friedrich, L., Furgeson, J., Greene, K., Kim, J., Lyskawa, J., Olson, C. B., & Smither Wulsin, C. (2016). *Teaching secondary students to write effectively* (NCEE 2017-4002). Washington, DC: National

Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education.

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, *99*, 445–476. http://dx.doi.org/10.1037/0022-0663.99.3.445

Gurney, D., Gersten, R., Dimino, J., & Carnine, D. (1990). Story grammar: Effective literature instruction for high school students with learning disabilities. *Journal of Learning Disabilities, 23*, 335-348.

Hernandez, D. J., Denton, N. A., & McCartney, S. E. (2008). Children in immigrant families: Looking to America's future. *Social Policy Report, 22*(3), 3-22.

Institute of Education Sciences (IES). (2017). *What Works Clearinghouse Standards Handbook, Version 4.0.* Washington, DC: U. S. Department of Education, What Works Clearinghouse.

Kanno, Y., & Cromley, J. (2015). English language learners' pathways to four-year colleges. *Teachers College Record*, 117 (120306), 1-44.

Kenward, M.G, & Roger, J.H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics, 53*, 983-997.

Kim, J., Olson, C. B., Scarcella, R., Kramer, J., Pearson, M., van Dyk, D., ... Land, R. (2011). A randomized experiment of cognitive strategies approach to text-based analytical writing for mainstreamed Latino English learners in Grades 6 to 12. *Journal of Research on Educational Effectiveness, 4*, 231-263. http://dx.doi.org/10.1080/19345747.2010.523513

Kiuhara, S. A., Graham, S., & Hawken, L. S. (2009). Teaching writing to high school

students: A national survey. *Journal of Educational Psychology, 101*(1), 136.

Klein, S. P., McArthur, D. J., & Stecher, B. M. (1995, February). What are the challenges to 'scaling up' reform? In Joining forces: Spreading strategies. Proceedings of the Invitational Conference on Systemic Reform. Washington, DC.

Langer, J.A. (1986). *Children reading and writing: Structures and strategies*. Norwood, NJ: Ablex.

Langer, J.A. (1991). Literacy and schooling: A sociocognitive perspective. In E.H. Hiebert (Ed.), *Literacy for a diverse society: Perspectives, practices, and policies* (pp. 10-27). New York, NY: Teachers College Press.

Langer, J. A. (2000). *Beating the odds: Teaching middle and high school students to read and write well* (Report No. 12014). Albany, NY: National Research Center on English Learning and Achievement.

Langer, J.A., & Applebee, A.N. (1986). Reading and writing instruction: Toward a theory of teaching and learning. In E.Z. Rothkopf (Ed.), *Review of research in education* (vol. 13, pp. 171-197). Washington, DC: American Educational Research Association.

MacArthur, C.A. (2019). Evaluation and revision. In S. Graham, C.A. MacArthur, & M. Hebert (Eds.), *Best practices in writing instruction* (3rd ed.), pp. 287-308. New York, NY: Guilford Press.

MacArthur, C.A., Philippakos, Z. A., & Ianetta, M. (2015). Self-regulated strategy instruction in college developmental writing. *Journal of Educational Psychology, 107*, 855-867.

Magnier, M. (2011, March 17). For one quake survivor, self-help in the face of seeming

helplessness. *The Los Angeles Times.* Retrieved from http://articles.latimes.com/2011/mar/17/world/la-fg-japan-quake-scuba-20110317

Maxwell, L. (2012). Raising Latino achievement seen as "demographic imperative." *Education Week, 31*, 4-5.

McLaughlin, M.C., & Mitra, D. (2001). Theory-based change and change-based theory; Going deeper and going broader. *Journal of Educational Change, 2* (4), 301-323.

Menken, K., Kleyn, T. (2009). The difficult road for long-term English learners. *Educational Leadership,* 66. (7). Retrieved from http://www.ascd.org/publications/educational_leadership/apr09/vol166/num07/TheThe_Difficult_Road_for_Long_Term_English_Learners.aspx

National Center for Education Statistics. (2017). *Percent of Latino Adults age 25 and older with a Bachelor's degree or higher in California, 2005 to 2015*. Retrieved from https://factfinder.census.gov

National Commission on Writing for America's Families, Schools, and Colleges. (2004, September). *Writing: A ticket to work...or a ticket out: A survey of business leaders.* College Board.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects.* Washington, DC: Author.

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for*

reading instruction (*NIH Publication No. 00–4769*). Washington, DC: U. S. Government Printing Office.

Noden, H.R. (2011). *Image grammar: Teaching grammar as part of the writing process.* Portsmouth, NH: Heinemann.

Nuñez, A. M., Rios - Aguilar, C., Kanno, Y., & Flores, S. M. (2016). English learners and their transition to postsecondary education. In M. B. Pauleson (Ed.), *Higher education: Handbook of theory and research* (pp. 41-90). New York: Springer.

Olsen, L. (2010). *Reparable Harm: Fulfilling the unkept promise of educational opportunity for California's long-term English learners.* Retrieved from https:www.californiastogether.org/?s=reparable+harm&submit=Search

Olson, C.B. (2011). *The reading writing connection: Strategies for teaching and learning in the secondary classroom* (3rd ed.). New York: Perason.

Olson, C. B., Kim, J. S., Scarcella, R., Kramer, J., Pearson, M., van Dyk, D., & Land, R. (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,* 323-355.

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,* 269-303.

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,* 259-269.

Olson, C. B., Matuchniak, T., Chung, H. Q., Stumpf, R. A., & Farkas, G. (2017). Reducing achievement gaps in academic writing for Latino secondary students and English learners. *Journal of Educational Psychology, 109*(1), 1-21. http://dx.doi.org/10.1037/edu0000095

Olson, C.B., Scarcella, R., & Matuchniak, T. (2015). *Helping English learners to write: Meeting Common Core Standards, grades 6-12.* New York: Teachers College Press.

Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology, 8*(3), 293-316.

Pearson, P. & Gallagher, M. (1983). The instruction of reading comprehension. *Contemporary Educational Psychology, 8,* 317-344. doi: 10.1016/0361-476X(83)90019-X

Pitts, L. (2010, January 14). Sometimes, the earth is cruel. *The Dallas Morning News.* Retrieved from http://www.dallasnews.com/opinion/latest-columns/20100114-Leonard-Pitts-Sometimes-the-9424.ece

Rosenblatt, R. (1982, January 25). The man in the water. *Time.* Retrieved from http://content.time.com/time/magazine/article/0,9171,925257,00.html

Sanchez, C. (2017, February 23). English language learners: How your state is doing. *NPR.* Retrieved from https://www.npr.org/sections/ed/2017/02/23/512451228/5-million-english-language-learners-a-vast-pool-of-talent-at-risk

Scardamalia, M., & Bereiter, C. (1987). Knowledge telling and knowledge transforming in written composition. *Advances in Applied Psycholinguistics, 2,* 142-175.

Schaalje, G.B., McBride, J.B., & Fellingham, G.W. (2002). Adequacy of approximations to distributions of test statistics in complex mixed linear models. *Journal of Agricultural, Biological and Environmental Statistics, 7*(4), 512-524.

Schochet, P. Z. (2008). *The Late Pretest Problem in Randomized Control Trials of Education Interventions* (NCEE 2009-4033). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Short, D., & Fitzsimmons, S. (2007). *Double the work: Challenges and solutions to acquiring language and academic literacy for adolescent English language learners-A report to Carnegie Corporation of New York.* Washington, DC: Alliance for Excellent Education.

Sternberg, R.J., Birney, D., Kirlik, A., Stemler, S., Jarvin, L., Grigorenko, E.L. (2006). From molehill to mountain: The process of scaling up educational interventions (firsthand experience upscaling the theory of successful intelligence). In Mark A. Costas & Robert J. Stenberg (Eds.). *Translating theory and research into educational practice: Developments in content domains, large-scale reform, and intellectual capacity* (pp. 205-222). Mahwah, N.J.: Lawrence Erlbaum.

Taylor, B., Pearson, D., Peterson, D., & Rodriguez, M. (2005). The CIERA School Change Framework: An evidence-based approach to professional development and school reading improvement. *Reading Research Quarterly, 40,* 40–69. http://dx.doi.org/10.1598/RRQ.40.1.3

The Education Trust-West. (2017). *The majority report: Supporting the educational*

success of Latino students in California. Retrieved from https://west.edtrust.org/resource/the-majority-report/

Tierney, R. J., & Pearson, P. D. (1983). Toward a composing model of reading. *Language Arts, 60*, 568-580.

U.S. Census Bureau, 2015 ACS 5-year estimates, "Sex by Age" and "Sex by Age Hispanic or Latino)," https:factfinder.census.gov.

U.S. Census Bureau. (2017). *Percent of youth under the age of 25 who are Latino, by county (2015)*. Retrieved from https://factfinder.census.gov

U.S. Department of Education, Institute of Educational Sciences, National Center for Education Statistics. (2012). *The nation's report card: Writing 2011* (NCES 2012-470). Washington, DC: Author.

U.S. Department of Education, National Center for Education Statistics. (2016). *The condition of education 2016* (NCES 2016-144), English Language Learners in Public Schools.

U.S. Department of Education, Institute of Educational Sciences, National Center for Education Statistics. (2017). *The condition of education: English language learners 2014-15*. Washington, DC: Author.

Vygotsky, L.S. (1986). *Thought and language*. Cambridge, MA: MIT Press.