Table of Contents

Project Narrative

A. Significance

1. Contribution to increased knowledge or understanding of educational problems

The demand for computer science (CS) knowledge and skills is increasing. The growth rate of U.S. computer and IT jobs in the U.S. is projected to outpace all other occupations, adding about 557,100 new jobs by 2026 (Bureau of Labor Statistics, 2018). And yet, if the current trend in CS degrees awarded continues, there will only be approximately 50,000 graduates to fill these job openings (Kena et al., 2015).

A critical task toward addressing this demand is broadening the CS pipeline. This is especially important for the large and growing Hispanic population -- which grew from 9 million (6% of U.S. population) in 1970 to 59 million (18% of population) in 2017, and is projected to reach 132 million (30% of population) by 2050, but is seriously underrepresented in CS education and achievement. For example, in California, Hispanics constitute 54% of the K-12 population, but only 22% of advanced placement CS test takers (College Board, 2017).

There are a number of important obstacles to effective CS instruction in schools, including lack of teacher expertise and insufficient time in the school day (Fancsali, 2018). Hispanic students face additional disadvantages in learning CS, including fewer course offerings in CS in Hispanic-neighborhood schools (Martin, McAlear, & Scott, 2015), reduced access to home computers or family members who are knowledgeable about CS (Royal & Swift, 2016), lack of Hispanic role models in CS whether through direct experience or through media representations (Wang et al., 2016), and decontextualized and individualized methods of CS instruction that are not a good match for the cultural values of Hispanic students and families, which tend to favor collaborative work that is meaningful for their communities (Brown &

Doolittle, 2008). For the large numbers of Hispanic students who are also non-native speakers of English, the abstract vocabulary of computing and strict syntactic demands of programming languages can pose additional challenges (Bennedsen & Caspersen, 2012). While there has been a growing body of research on two related fields, English learners in STEM and CS for all, there is almost no research on English learners in CS, thus providing no roadmap for educators as to how to improve CS knowledge, skills, and attitudes for this population.

While there are a number of promising initiatives geared toward improving access to quality CS education in high school (e.g., Goode & Margolis, 2011), by that age, wide disparities already exist in diverse groups' knowledge, skills, and attitudes toward STEM (Carlone, Scott, & Lowder, 2014), making it difficult for those on the margins to catch up (Shettle et al., 2007). This is especially true in CS, since low-SES learners typically lack access to the kinds of out-of-school experiences through robotics clubs, coding summer camps, and at-home mentoring that help promote early CS knowledge and identity (Barron et al., 2009).

Finally, access to quality CS education for English learners, especially in elementary school programs, is highly complex and challenging. First, the elementary school curriculum is already packed and test-driven, incentivizing teachers to prioritize tested subjects, such as English and math, rather than non-tested subjects, such as CS. Additionally, only a minority of elementary school teachers themselves have the CS content and pedagogical knowledge to confidently introduce the subject to their students.

This project will develop and evaluate an integrated curriculum and professional development intervention focused on the computational thinking that is foundational to CS (Wing, 2006) and the CS identity of Hispanic students in fourth grade. Our proposal tackles some of the biggest barriers to CS attainment by Hispanic students: (a) insufficient instructional time during the school day, (b) lack of teacher expertise, (c) lack of role models, and (d) lack of scaffolding for non-native speakers of English. It will illuminate the types of scaffolding that are most valuable for CS attainment by Hispanic students by answering the following research question: Can this intervention, integrated within English language arts (ELA) instruction and with language and learning scaffolding, result in significant gains in computational thinking and CS identity while maintaining students' levels of English language and math proficiency?

2. Development or demonstration of promising new strategies

The starting point for the project is the Creative Computing Curriculum Guide developed at the Harvard Graduate School of Education. This curriculum, first released in 2011 and based on the Scratch visual programming language, is one the most widely used in the country for elementary school programming. The intervention that will be iteratively developed and evaluated in this project is based on three promising innovations to this curriculum, each developed by different teams within the project but not yet combined. Together, we believe they will greatly amplify the potential of this curriculum to be successfully used in elementary schools with high percentages of Hispanics, English learners, and other high-need groups.

Integration into ELA. Elementary schools with high percentages of English language learners (ELLs) devote large amounts of instructional time to improving students' English skills. This makes it challenging to introduce non-core curriculum, such as CS. Indeed, research has shown that even science, let alone CS, is rare in high-ELL schools and districts (Gomez-Zwiep, 2017). San Francisco Unified School District (SFUSD) has addressed this challenge by adapting the Creative Computing Curriculum for integration into ELA instruction. The curriculum exploits the affordances of Scratch for learning to decode and code stories of the same genres that are emphasized in elementary school. It also integrates age-appropriate readings about diverse leaders in CS, thus strengthening the connection to reading while also providing culturally relevant support.

The SFUSD curriculum was evaluated in a study by the University of Chicago team, which found that students of all levels who completed the curriculum demonstrated knowledge of key computational thinking concepts, but with gaps between the performance of students in low-performing and high-performing schools (Salac et al., 2019). To close those gaps, the partners have been developing additional linguistic and learning scaffolding as discussed below.

Linguistic scaffolding. A team at the University of California, Irvine (UCI), in partnership with Santa Ana Unified School District (SAUSD), has been working with the SFUSD curriculum to develop additional language scaffolding to amplify its effectiveness with English learners, following best practices recommended by a national panel (National Academies of Science, Engineering, and Medicine, 2018). First, the revised curriculum will integrate CS and ELA tasks to **engage students in disciplinary practices** through collaborative exploration, modification, and creation of products, providing authentic contexts for language use while making instruction more engaging for multilingual students (Janzen, 2008; National Research Council [NRC], 2006, 2012; Rosebery & Warren, 2008). Computer science disciplinary activities and learning goals are aligned with standards so as to best guide teachers (see Table 1 below for an example).

Second, the revised curriculum will **encourage rich classroom discourse** through explicit suggestions of activity formats (e.g., individual thinking time, pair programming, small group, whole class) that encourage students to use disciplinary language in multiple contexts. The professional development associated with the curriculum will focus on teacher noticing of students’ discourse to facilitate productive talk (Shea & Shanahan, 2011).

Table 1. Sample learning goals with Grade 4 Common Core ELA, ELD, and CSTA Standards

<table>
<tr><th colspan="3">Activity: Students program a story about their lives, families, or communities</th></tr>
<tr><th colspan="3">Computer Science Concepts: Loops, Sequences, Conditionals</th></tr>
<tr><td colspan="3">Computer Science Teacher Association (CSTA) Standards</td></tr>
<tr><td>CSTA 1B-AP-10</td><td colspan="2">Create programs that include sequences, events, loops, and conditionals</td></tr>
<tr><td>CSTA 1B-AP-13</td><td colspan="2">Use an iterative process to plan the development of a program by including others’ perspectives and considering user preferences</td></tr>
<tr><td>CSTA 1B-AP-15</td><td colspan="2">Test and debug a program or algorithm to ensure it runs as intended</td></tr>
<tr><td colspan="3">English Language Development (ELD) Standards</td></tr>
<tr><td>Emerging</td><td>Expanding</td><td>Bridging</td></tr>
<tr><td>3. Offering opinions
Negotiate with or persuade others in conversations using basic learned phrases (e.g., I think) as well as open responses in order to gain and/or hold the floor.</td><td>3. Offering opinions
Negotiate with or persuade others in conversations using a variety of learned phrases (e.g., That’s a good idea. However …) as well as open responses, in order to gain and/or hold the floor.</td><td>3. Offering opinions
Negotiate with or persuade others in conversations using a variety of learned phrases (e.g., That’s a good idea. However …) as well as open responses in order to gain and/or hold the floor, elaborate on an idea, and provide different opinions.</td></tr>
<tr><td>11. Supporting opinions
Offer opinions and provide good reasons (e.g., My favorite book is X because X) referring to the text or to relevant background knowledge.</td><td>11. Supporting opinions
Offer opinions and provide good reasons and some textual evidence or relevant background knowledge (e.g., paraphrased examples from text or knowledge of content).</td><td>11. Supporting opinions
Offer opinions and provide good reasons with detailed textual evidence or relevant background knowledge (e.g., specific examples from text or knowledge of content).</td></tr>
<tr><td colspan="3">Corresponding English Language Arts Standards</td></tr>
<tr><td>CCSS.ELA-L.SL.4.1</td><td colspan="2">Engage effectively in a range of collaborative discussions with diverse partners, building on others’ ideas and expressing their own clearly.</td></tr>
<tr><td>CCSS.ELA-L.SL.4.4</td><td colspan="2">Report on a topic or text, tell a story, or recount an experience in an organized manner, using appropriate facts and relevant, descriptive details to support main ideas or themes; speak clearly at an understandable pace.</td></tr>
<tr><td>CCSS.ELA-L.SL.4.6</td><td colspan="2">Differentiate between contexts that call for formal English (e.g., presenting ideas) and situations where informal discourse is appropriate (e.g., small-group discussion); use formal English when appropriate to task and situation.</td></tr>
<tr><td>CCSS.ELA-L.W.4.9</td><td colspan="2">Draw evidence from literary or informational texts to support analysis, reflection, and research.</td></tr>
</table>

Third, strategies that teachers can use to **build on students’ existing resources** to acquire proficiency in language and CS will be highlighted in the curriculum and during professional development, including tips for teacher “talk moves” (Michaels & O’Connor, 2015) such as asking for clarification and leveraging students’ own ways of explaining to guide them towards more formal language and advanced CS concepts.

Fourth, visualizations and physical, unplugged activities will be built into the curriculum to **engage students in multiple modalities,** including linguistic modalities of talk and text, as well as nonlinguistic modalities such as gestures, pictures, and symbols, to better teach key academic vocabulary and computational thinking concepts (cf., Lee, Llosa, Grapin, Haas, & Goggins, 2019). For example, students will learn about the concept and term, "parallelism," first through an unplugged activity encompassing body movement and then through visualizations.

Fifth, the curriculum will **provide explicit focus on how language functions in the discipline** by providing language frames to teachers for use by students during peer feedback and pair programming, and while asking for assistance (see example in Table 2 below).

Table 2. Computer Science Language Functions

Teacher Activities	Student Discourse			CS Concepts (Language Function)
	Emerging	**Expanding**	**Bridging**	
Remind students to think about the events that will cause each action to happen in their project, which programs will run parallel to each other, and how their project will reset once it has finished running.	I need help with ___. ___ caused ___to happen. ___ and ___ are running at the same time. I used ___ to reset the program.	I am having difficulty with ___. __ is the event that caused __ to happen. __ and __ are running parallel to each other. I used _ to initialize the program.	Could you help me fix the following challenge in my code ___? The event that caused ___ to happen is ___. _ and _ are running parallel to each other/simultaneously/ at the same time. __ caused the program to initialize.	Debugging, events, initialization, parallelism (Describing, comparing)

Observations, interviews, teacher design meetings, and pre-post survey data collected by the UCI team have demonstrated that the revised curriculum with linguistic scaffolding is feasible for implementation, successfully engages students of diverse language backgrounds, and increases Hispanic students' identity with CS (Jacob et al., 2018; Jacob et al., 2019).

CS learning scaffolding. Inquiry-based instruction enables learners to work on motivating projects. However, placing students into unstructured inquiry without sufficient preparation can lead to frustration and wasted opportunities (Bransford et al., 2009). To best achieve the advantages of inquiry-based instruction, this project will modify the SFUSD curriculum to incorporate a scaffolded CS instructional approach known as Use-Modify-Create (Lee et al., 2011; Werner et al., 2012) in which students will first *use* existing programs, then work together to *modify* them, and finally *create* their own. As a further innovation, we will incorporate an extra layer of scaffolding within the "Use" stage -- a learning strategy known as "TIPP&SEE" developed by the Computing for ANyONe (CANON) lab at the University of Chicago and faculty at Texas State University. Based on the reading comprehension strategy THIEVES (Manz, 2002), TIPP&SEE is designed to focus students on the context of the project so they are ready to learn the intended material. Students are taught to first examine the *Title* of the program, then analyze its *Instructions*, and consider the *Purpose* of the program to think about what they will learn through using it. Then, they are taught how to find what code controls the actions in a Scratch program, providing a roadmap for navigating the complex user interface. Next, they *Play* with the program to test its features, recording what they observe. Students then transition to finding the code of the program, choosing a *Sprite* within the program, and seeking out the *Event* that caused the sprite to perform an action. Finally, the strategy scaffolds the process of how to learn from provided code. Students *Explore* code with provided prompts, making different kinds of changes (e.g., changing a number in an instruction, removing an instruction, replicating an instruction) and observing how those changes affect how the program runs. This process allows students to learn through the Use stage (of the Use-Modify-Create approach) so that they can be adequately prepared to Modify the program.

Use-Modify-Create and TIPP&SEE have been integrated into two CS curricula and piloted in Chicago Public Schools (CPS) and the Austin Independent School District. Teachers and researcher-observers from the University of Chicago and Texas State University have reported high levels of engagement and learning among Hispanic students.

This proposal meets **Absolute Priority 1** (Demonstrates a Rationale), **Absolute Priority 3** (Field-Initiated Innovations Promoting STEM Education, with a Particular Focus on Computer Science), and **Competitive Preference Priority** (Project Designed to Improve Student Achievement or other Educational Outcomes in Computer Science).

Rationale. This project expands access to and participation in rigorous CS coursework for traditionally underrepresented students, specifically Hispanic students. The three elements of the curriculum, integration into ELA, linguistic scaffolding, and CS learning scaffolding, are all based on prior high-quality research.

Integration into ELA. Research suggests that CS is not systematically integrated into elementary school curriculum (Israel et al., 2015). Students' experiences with CS are often one-off (such as an "hour of code") or extra-curricular, with Hispanics, English learners, and other high-need youth the least likely to receive systematic CS instruction (Martin et al., 2015). Though elementary teachers often seek to integrate computing into math and science (Weintrop et al., 2016), the foundations of coding are more synergistic with ELA, since initial programming fits well with the narrative and informative genres taught in elementary school (Jacob & Warschauer, 2018). A review by Burke and Kafai (2012) found that programming stories in the narrative genres fosters literacy development, increases technological fluency, and promotes coding at an early age. Similarly, a study by de Souza et al. (2011) showed that discursive structures typically found in students' narratives serve as a heuristic to understand the process underlying computational thinking. A case study by Peppler and Warschauer (2012) illuminated several interlocking features of coding and literacy, drawing young children's attention to symbol-meaning relationships, providing multimodal scaffolding for learning letters and words, and offering a highly engaging and supportive environment for low-literacy children to demonstrate their skills and abilities. Opportunities for creativity and self-expression in coding have been found to motivate students who might not ordinarily view themselves as computer programmers (Kelleher & Pausch, 2007; Peppler & Kafai, 2007), facilitating the kinds of open-endedness, links to students' culture and community, and opportunity to be recognized as experts that increase positive identity development in STEM (NRC, 2014).

*Linguistic scaffolding***.** Integrating disciplinary practices and linguistic scaffolding into the computational thinking curriculum is based on well-established findings on how to best engage English learners in STEM practices (National Academies of Sciences, Engineering, and Medicine, 2018). As **students engage in STEM disciplinary practices**, they negotiate ideas with peers to co-construct meaning in STEM communities of practice. The discipline of CS requires students to grapple with abstract, decontextualized language, and leverage prior knowledge as they move from concrete to more abstract concepts. Engaging students in disciplinary practices through inquiry-based instruction significantly increases gains on science achievement for elementary English learners (Estrella, et al., 2018). Given its efficacy in STEM subjects, this approach holds great promise for the field of CS.

Learners construct STEM knowledge through scientific discourse (Kelly & Chen, 1999; McGinn & Roth, 1999). Collaborative activities that **encourage rich classroom discourse** such as pair programming, group work, and peer feedback facilitate opportunities for language and content area growth. Shea et al. (2017) found that integrating language and content in a manner that facilitated peer-to-peer talk increased both ELA and math scores for ELLs. Furthermore, inquiry-based approaches **build on students' existing resources** by allowing students to use their everyday sense-making abilities to access content (Brown & Ryoo, 2008). In addition to collaboration and inquiry, unplugged activities, visualizations, and scaffolding strategies **engage students in multiple modalities** to better teach key language and concepts, as evidenced by Ryoo et al.'s (2018) study on the value of visualizations for helping English learners understand science. Finally, linguistic scaffolds will **provide explicit focus on how language functions in the discipline**. In CS, these functions include discussing the relationship between desired outcomes and coding strategies, and analyzing errors in coding that prevent those outcomes from being achieved (Jacob et al., 2018). Explicit teaching of the corresponding language forms and functions reinforces students' understanding while developing their linguistic repertoires. Elementary students in a program that integrated content, ELD standards, and corresponding linguistic frames into inquiry-based, STEM unit plans showed significant increases on ELD and ELA scores compared to students in a traditional program (Zwiep & Straits, 2013).

*CS learning scaffolding and strategy instruction***.** The rationale for CS learning scaffolding via strategy instruction is based upon years of research in reading comprehension and literacy. Good readers have a purpose for reading and apply a variety of strategies as they read to help them construct meaning from the text (Pressley, 2002). Further, they are able to adapt their strategy use based on the type of text, the difficulty of the text, and their reading purpose. Before reading, they set goals and consider organization and text structure (Schunk, 2003). During reading, they self-monitor and self-question, consider relationships between ideas, and make connections to prior knowledge (Pressley, 2002). In comparison, poor readers are non-strategic in their approach and do not engage in the metacognitive activities that allow them to self-monitor and achieve comprehension. Explicit teaching of reading comprehension strategies has been found to help students understand what skilled readers are doing, support them in thinking about their own thinking, and proceduralize the steps to accomplish their comprehension goals (Palincsar, 1986). An experimental study found that students with learning disabilities who received reading strategy instruction overcame the gap in reading comprehension with their normally achieving peers (Johnson et al., 1997). Similarly, a meta-analysis of learning strategy instruction revealed it was valuable in all cases and at all levels (Donker et al., 2014). The Use-Modify-Create pedagogical approach (Lee et al., 2011) and TIPP&SEE strategies build from these lessons by providing students the structured and scaffolded learning strategies that have been effective in broader domains of learning.

B. Quality of the Project Design

1. Clearly specified and measurable goals, objectives, and outcomes

The project will be carried out in two phases, research and development (R&D; Years 1-3, discussed further below), and evaluation (Years 4-5, discussed further in Section D). We will iteratively develop and evaluate both a yearlong curriculum and a 5-day teacher professional development (PD) program. See Appendix I-1 for sample curriculum and Appendix I-2 for a sample PD agenda.

R&D Stage. The first three years of the project will be considered the R&D stage. The curriculum, professional development, and measures will be developed and refined in Year 1. In Year 2 and (following iterative revision) Year 3, they will be piloted in six schools, two each in SFUSD, CPS, and SAUSD (see Letters of Commitment in Appendix C). These six schools will be purposely chosen to provide iterative data on the implementation of the intervention and measures with the main targeted population (Hispanic ELLs) as well as additional disadvantaged populations (African Americans, low-income learners, Hispanics who are not ELLs, ELLs who are not Hispanics). In addition, the outcome measures (see Section C. 3, below) will be implemented in one control school in SAUSD in Year 3 for purposes of assisting their validation.

Table 3. Demographics of participating districts

District	Free/Reduced Lunch	Hispanic	English Learners
San Francisco Unified (SFUSD)*	52.1%	31.2%	28%
Chicago Public Schools (CPS)**	76.6%	46.7%	18.7%
Santa Ana Unified (SAUSD)*	80.4%	92.9%	38.7%

*Source: *www.ed-data.org; **https://cps.edu/About_CPS/*

The following implementation data will be used to assess how the curriculum is being used in diverse classrooms and its impact on particular groups of students to inform further iterations:

Classroom observations. We will conduct quarterly classroom observations to provide teacher support, ensure adequate communication between teachers and the research team, view implementation in a variety of settings, and identify any struggles teachers or students are facing as well as promising strategies that can be shared through subsequent teacher development.

*Teacher and administrator design groups***.** Teachers implementing the curriculum will meet monthly during the school year for professional development where they will discuss the strengths and weaknesses of the curriculum and the range of ways it is being implemented. During design groups, researchers will take detailed notes on teacher and administrator feedback and collaboratively develop action plans to address practitioner insights. The project's independent evaluator, WestEd, will be invited to attend the final meeting of each year, where the agenda will focus on best practices for implementation, areas for improvement, and overall feedback on the intervention.

Table 4 below shows the project's specific, measurable goals, objectives, and outcomes.

Table 4. Goals, objectives, and outcomes

Goal 1	Develop, pilot, and implement CS curriculum integrated into ELA instruction with linguistic scaffolds and CS learning scaffolds
1.1	By August 2020, develop a yearlong curriculum that is integrated into ELA instruction with linguistic scaffolds and CS learning scaffolds. **Measures:** Lesson plans, student workbook.
1.2	By the end of the 2020-21 school year, implement new curriculum with students in one grade at 2 elementary schools each in SFUSD, CPS, and SAUSD. **Measures:** Online Scratch project studios for each class; teacher surveys; minutes of design group meetings.
1.3	By the end of the 2021-22 school year, implement a *revised* curriculum with students in one grade at 2 elementary schools each in SFUSD, CPS, and SAUSD. **Measures:** Online Scratch project studios for each class; teacher surveys; minutes of design group meetings.
1.4	In the 2022-23 school year, randomly assign half of 24 schools in SAUSD *that have not yet implemented the curriculum* to treatment in RCT Year 1 and implement the curriculum in all fourth-grade treatment classes. Administer outcome measures to students in both treatment and control schools. **Measures:** Online Scratch project studios for each class; teacher surveys; minutes of design group meetings; completed outcome measures.
1.5	In the 2023-24 school year, implement the curriculum in the 12 treatment schools from the prior year and the 12 delayed-treatment schools. **Measure:** Online Scratch project studios for each class.
1.6	By the end of the 2023-24 school year, at least 4,320 students (at least 75% of whom are Hispanic students) will have been enrolled in classes implementing the curriculum. **Measures:** Online Scratch project studios for each class; teacher surveys; class records from district.
Goal 2	Develop and implement professional development for teachers related to new curriculum
2.1	By August 2020, develop professional development for the new curriculum. **Measure:** Program agenda.
2.2	By September 2020, provide professional development for the participating elementary teachers from 2 schools each in SFUSD, CPS, and SAUSD. **Measure:** Attendance logs.
2.3	By September 2021, provide refresher short professional development for the participating elementary teachers from 2 schools each in SFUSD, CPS, and SAUSD. **Measure:** Attendance logs.
2.4	In the 2022-23 school year, provide professional development for the participating teachers in the 12 schools in the RCT treatment condition. **Measure:** Attendance logs.
2.5	By the end of the 2023-24 school year, at least 90 total teachers will have received professional development in implementing the curriculum. **Measures:** Attendance logs.
Goal 3	Validate outcome measures of computational thinking and CS identity
3.1	By the end of October 2020 and October 2021, implement *Lean Computational Thinking Abilities Assessment* (LCTAA) and *Is Computer Science Me?* (ICSM) survey with students at 2 schools each in SFUSD, CPS, and SAUSD and 1 control class in Year 3. **Measures:** Completed tests and surveys.
3.2	By the end of the 2020-21 and the 2021-22 school years, implement LCTAA and ICSM with students at 2 schools each in SFUSD, CPS, and SAUSD and 1 control class in Year 3. **Measures:** Completed tests and surveys.
3.3	Summer 2021 and 2022, calculate reliability and other psychometric properties of pre- and post-test LCTAA and ICSM measures. **Measures:** Cronbach's alpha, Item and Test Characteristic

	Curves, Infit and Outfit metrics of Rasch Model.
Goal 4	Obtain evidence of intervention's effectiveness
4.1	By the end of the 2022-23 academic year, at least 80% of teachers in the RCT treatment condition will have implemented 75% of the new curriculum. **Measures:** Online Scratch project studios for each class; teacher surveys.
4.2	By the end of the 2022-23 academic year, students in treatment schools will show statistically significant improvement with a moderate effect size (0.35-0.45) on their LCTAA and ICSM scores, and similar annual ELA and math NWEA MAP or SBAC assessment scores as control peers. **Measures:** Completed tests and surveys; annual ELA and Math NWEA MAP and SBAC assessments.
Goal 5	Disseminate results; secure continuation of project funding
5.1	Prepare and submit for publication at least one article on results or findings to date each year of the project after Year 1. **Measure:** Proof of submission.
5.2	Make at least one presentation to an academic or practitioner audience on the project each year after Year 1. **Measure:** Program agenda or acceptance.
5.3	Prepare and submit a grant proposal for continuation of the project to NSF, IES, or other funder in each of Years 3, 4, and 5. **Measure**: Filing verification from agency/funder.

2. A conceptual framework underlying the proposed research

Universal Design for Learning (UDL) provides our conceptual framework, particularly its three foundational principles: (a) multiple means of representation -- give learners various ways of acquiring information and knowledge; (b) multiple means of expression -- provide learners alternatives for demonstrating what they know; and (c) multiple means of engagement -- tap into learners' interests, offer appropriate challenges, and increase motivation. A central theme of UDL is that if these multiple means are "baked into" the design of the learning environment from the very beginning, it will make learning more accessible to both the specifically targeted group as well as other related groups. The curriculum will be designed with the particular needs of Hispanic ELLs in mind. However, we believe that this design will also make the curriculum more accessible to other at-risk groups, including Hispanics who are not ELLs, ELLs from other linguistic backgrounds, and low-income learners regardless of the language or ethnic background.

In this project, multiple means of representation will be reflected in the visualizations and unplugged activities that convey important concepts. Multiple means of expression will be supported through opportunities for rich discourse in pairs and small groups and students creating their own programs on relevant topics. Multiple means of engagement will be provided through opportunities to read stories, watch videos, and code as individuals and in pairs on topics relevant to their interest, while the Use-Modify-Create and TIPP&SEE strategies will ensure that students have appropriate levels of challenge so as to increase their motivation.

3. Ensuring feedback and continuous improvement in the operation of the project

In the R&D stage, we will utilize the Plan/Do/Study/Act (PDSA) framework for continuous improvement in discipline practices. Each PDSA cycle begins with articulating the change and recording predictions about what we expect will happen (plan); attempting the change and documenting what in fact did happen (do); comparing the results to the predictions (study); and then deciding on what to do next (act; Grunow, 2015). In many instances, although a PDSA cycle may not generate the results expected, it provides clues as to what to try instead and becomes the basis for the next PDSA cycle (Grunow, 2015). During the “do” cycle, we will collect data to inform the next stage, “study,” including our outcome measures and implementation measures. The Leadership Team and External Evaluator will then review the results (study) and determine what, if any, modifications are needed to better reach our planned change at the teacher and student level. If modifications are needed, we will implement them (act) and share results of our research and evaluation findings through conferences, presentations, articles, and other dissemination activities. We will repeat the cycle to ensure continuous improvement until the end of the project. To refine our intervention, we will use Year 5 to take a closer look at teachers with varying degrees of success in the RCT and implement any suggested changes to the curriculum and PD for final dissemination.

Dissemination. Our dissemination activities will provide additional opportunities for feedback. We will reach out to **district stakeholders in SAUSD, SFUSD, and CPS** through presentations to principals' meetings and the Board of Education, newsletters to parents in the participating schools, and showcase events where relatives and friends can come view students' Scratch projects. SFUSD will incorporate the revised materials into its current elementary CS curriculum, thus reaching an additional 8,000 elementary school students a year. We will reach out to **educators across the country** through presentations and workshops at practitioner-oriented conferences and journals including those hosted by the ACM Special Interest Group on Computer Science Education (SIGCSE), the Computer Science Teacher Association (CSTA), Computer Using Educators (CUE), EdSurge Fusion, and Digital Promise. We will reach out to **scholars** through presentations and publications at appropriate academic venues such as the ACM SIGCSE Technical Symposium; Research on Equity & Sustained Participation in Engineering, Computing, & Technology (RESPECT); International Computing Education Research; and the American Educational Research Association (AERA), and papers in peer-reviewed journals such as *Computer Science Education*, *Computers & Education*, and *TESOL Quarterly*. To reach **policy makers**, Co-PI Richardson will share information through the Alliance for California Computing Education for Students and Schools (ACCESS), the state's leading advocate for CS education in California and a group that she previously chaired. Finally, detailed project information on PI Warschauer's Digital Learning Lab website and Co-PI Franklin's CANON website will also reach all of the above audiences.

C. Adequacy of the Resources and Quality of the Management Plan

1. The management plan is adequate to achieve the objectives of the project

The timelines and milestones for accomplishing project tasks are designed with clearly defined responsibilities to support our plan for an on-time, within-budget project. The timeline is summarized below in Table 5; for details, see Table 4, Appendix I-4 (Table 7), and Appendix I-3 for the complete Management Plan.

Table 5. Timeline

	Year 1				**Year 2**				**Year 3**				**Year 4**				**Year 5**			
	Oct 2019-Sept 2020				Oct 2020-Sept 2021				Oct 2021-Sept 2022				Oct 2022-Sept 2023				Oct 2023-Sept 2024			
	Develop				Pilot 1.0				Pilot 2.0				RCT				Dissemination			
					6 schools				6 schools				12/24 schools				24/24 schools			
	F	**W**	**S**	**Su**	**F**	**W**	**S**	**Su**	**F**	**W**	**S**	**Su**	**F**	**W**	**S**	**Su**	**F**	**W**	**S**	**Su**
Curriculum																				
Develop/Revise	■	■	■	■	■	■	■	■	■	■	■	■								
Implement					■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Teacher PD																				
Develop/Revise	■	■	■	■	■	■	■	■	■	■	■	■								
Implement				■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	
Measures																				
Develop/Revise	■	■	■	■				■				■								
Administer					■		■		■		■		■		■		■		■	
Validate								■				■								
Analysis																				
Implementation data					■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Outcome data																	■	■	■	■
Dissemination					■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Partnership Mgmt																				
MOUs with SFUSD, CPS, and SAUSD		■				■				■				■				■		
Match contributions confirmed			■				■				■				■				■	
Convene project teams	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■

2. The qualifications of key project personnel

The project's key personnel are highly qualified and experienced in all necessary areas covered by the project. An overview of project staff is provided below in Table 6, with resumes and outline of required qualifications in Appendix B.

Table 6. Personnel

<table>
<tr><th colspan="3">Leadership Team</th></tr>
<tr><td>Dr. Mark Warschauer, PI
Professor, Education & Informatics, UC Irvine</td><td colspan="2">Oversee all curriculum and professional development and direct the portions related to linguistic scaffolding for ELLs; oversee the pilot (in SAUSD and SFUSD) and RCT testing (SAUSD); coordinate revision of measures; ensure adherence to timelines, budgets, milestones; oversee Plan/Do/Study/Act activities; manage partnerships and communication.</td></tr>
<tr><td>Dr. Diana Franklin, Co-PI
Research Associate Professor, Computer Science, UChicago</td><td colspan="2">Direct curriculum and professional development related to learning scaffolding and measure refinement; oversee the pilot testing at CPS.</td></tr>
<tr><td>Dr. Debra Richardson, Co-PI
Professor, Informatics, UC Irvine</td><td colspan="2">Help ensure match of the curriculum with Computer Science Content Standards; assist with dissemination of the research within the computer science education community.</td></tr>
<tr><th colspan="3">Implementation Partners</th></tr>
<tr><td colspan="3">Personnel from each of our school districts will be involved in revising the curriculum and creating teacher development materials to ensure that the intervention meets the needs of each district; providing feedback on teacher PDs and lesson design; coordinating participation in the pilot testing and RCT; and assisting with dissemination of our findings.</td></tr>
<tr><th>SFUSD</th><th>CPS</th><th>SAUSD</th></tr>
<tr><td>Bryan Twarek, District Computer Science Supervisor</td><td>Andy Rasmussen, Computer Science Project Developer</td><td>Daniel Allen, Assistant Superintendent of Teaching & Learning</td></tr>
<tr><td>Bill Marsland, Elementary Computer Coordinator</td><td></td><td>Bianca Barquin, Director of Elementary Curriculum & Instruction</td></tr>
<tr><td></td><td></td><td>Don Isbell, Director of Career Technical Education</td></tr>
<tr><th colspan="3">External Evaluators</th></tr>
<tr><td>Yvonne Kao, Senior Research Associate, WestEd</td><td colspan="2">Oversee all major evaluation activities and manage the subcontract; direct pilot testing for the LCTAA and ICSM survey; direct quantitative data analysis for WestEd.</td></tr>
<tr><td>Mingyu Feng, Senior Research Associate, WestEd</td><td colspan="2">Project manager for the RCT, overseeing all aspects of data collection and reporting.</td></tr>
</table>

3. The potential for continued support of the project

We believe that the potential for continued support of the project after Federal funding ends is strong and the project team is committed to soliciting such funding prior to the end of the proposed project. We have built into the Management Plan the identification of funding opportunities and solicitation of grants beginning Summer 2022 (identification) and Summer

2023 (applications). Warschauer and Franklin have a strong history of funded research projects and are committed to further evaluating this innovative curriculum.

D. Quality of the Project Evaluation

1. Methods of evaluation to produce evidence meeting WWC standards

The evaluation will be carried out by WestEd. In Year 4, we will conduct a delayed-treatment RCT designed to meet WWC standards without qualification. The RCT will be carried out in 24 schools in SAUSD that have never implemented the curriculum. Twelve schools (approximately 1,080 participating students) will be randomly assigned to implement the intervention in fourth grade in Years 4 and 5, and the remaining 12 schools will implement business-as-usual ELA curriculum in fourth grade in Year 4 and the intervention in Year 5 (an additional 2,160 students). All 24 schools will collect pre- and post-data in Year 4.

2. Providing guidance on strategies for replication and testing in other settings

The evaluation will provide guidance about effective strategies suitable for replication or testing in other settings. The curriculum and professional development will be pilot tested in 6 schools, in 3 different districts, over two years. The pilot schools will be selected so that one in each district has a large population of Hispanic ELLs, our target population, and one in each district serves an additional diverse population (African Americans in CPS; ELLs in SFUSD; and a largely Hispanic school but with fewer English learners in SAUSD). This design enables robust understanding of implementation variation across districts and populations and provides time for the project team to test, revise, and re-test the curriculum, professional development, and measures to ensure they are fully developed prior to RCT. The RCT will take place in 24 schools. The RCT, together with the implementation studies, will help the team document differences in implementation to inform replication in various school settings, in terms of whether and how to implement the curriculum and the supports needed to help implement the curriculum with integrity. Our moderator analyses will also provide insights on the differential impacts of the curriculum across school settings and student populations. The qualitative and quantitative analyses of implementation data collected through observations (both during the implementation and the targeted focal teachers in Year 5), interviews, design group minutes, and surveys will help identify challenges in implementation, perceived educator and student progress/engagement, areas for improvement, and best practices utilized and whether they are suitable, sustainable, and replicable for different school settings.

3. Providing valid and reliable performance data on relevant outcomes

Evaluation activities will begin in Years 2-3 with validation of two proximal student outcome measures: the *Lean Computational Thinking Abilities Assessment* and the *Is Computer Science Me?* survey.

Lean Computational Thinking Abilities Assessment (LCTAA). The first measure of computational thinking abilities for K-12 (Wiebe et al., 2019) drew questions from two assessments, Computational Thinking Test and Bebras. These two assessments were designed to measure computational thinking outside the context of particular programming languages and can be administered to students without any coding experience, providing a way to compare computational thinking among students who have and have not received a programming curriculum (Saez-Lopez et al., 2016). The LCTAA consists of 25 items from these two assessments that were fully validated (see Appendix I-5 for examples of the questions). We will draw on a subset of these 25 questions most appropriate for the age level of our intervention. In addition, we will test additional questions developed by Dr. Franklin’s team in Years 2 and 3 to ensure that we have sufficient items to reliably assess the relevant computational skills and that the questions are age-appropriate for our population. The assessment will be piloted in Years 2-3. During the pilot years, WestEd will combine matrix sampling methods (Childs & Jaciw, 2003) with true score test equating (von Davier & Wilson, 2007) from a single-parameter item-response model (also known as a Rasch model). This method of psychometric analysis allows us to determine item and test characteristics on the whole pool of items (Wright, 1997), confirm the nature of the computational thinking construct being measured (de Ayala, 2010), and draw preliminary, classroom-level conclusions regarding the efficacy of the piloted intervention whilst being considerate of testing time (Childs & Jaciw, 2003). Using these calibrated items allows us to construct a proximal measure of the intervention efficacy that remains a reliable metric of students while being manageable in light of time or testing constraints.

Pre- and post-survey results on attitudes and beliefs towards CS. Items from the validated survey *Is Science Me?* (ISM; Gilmartin et al., 2006; see Appendix I-6) were adapted to capture students' attitudes towards CS disciplines and careers and the influence that families and peers have on student identification with computing (adapted survey renamed *Is Computer Science Me?* [ICSM]). In a prior study, we examined the internal consistency of the pre-test survey items under the original constructs in the ISM survey and found moderate internal consistency with Cronbach's alphas ranging from .56 to .68. The items will be revised by the team in an effort to improve reliability, and the revised survey will be piloted in Years 2-3. As with the LCTAA, each year WestEd will use a Rasch model (this time using the rating scale model variant; de Ayala, 2010) to identify best-functioning items to maximize reliability.

WestEd will also collect distal outcome measures in the form of standardized tests in ELA and mathematics. Each spring, students in California complete the Smarter Balanced Assessment Consortium (SBAC) test, which is aligned to the Common Core State Standards. For the 2016-2017 school year, the most recent for which data are available, SBAC reports an overall marginal reliability of 0.9 for 4th-grade ELA and 0.93 for 4th-grade mathematics (SBAC, n.d.). Students in SAUSD also take the NWEA MAP Growth assessments in reading and math. These assessments are also aligned to the Common Core State Standards. MAP Growth reading scores accurately predict students' SBAC proficiency levels 84% of the time for ELA and 88% of the time for math (NWEA Psychometrics Service Team, 2017). Students in SAUSD take the MAP Growth at three time points during the year: in August/September, December/ January, and March/April. Data from the first and third administrations will be used in our impact analyses.

Statistical power. We used Optimal Design (Raudenbush et al., 2011) to conduct power analyses for a three-level cluster-randomized trial with students nested within teachers within schools and measuring student-level outcomes. After attrition, we assumed three teachers would participate in each school, with an average of 25 students per teacher. Because few RCTs have been conducted with computational thinking interventions, we base our power analysis on Hedges and Hedberg (2007), who reported estimates of school-level intra-class correlations (ICCs, ρ_3) and the proportion of variance explained by pre-test and demographic covariates (R^2) on math and ELA standardized tests, both of which are distal outcomes for this study. The power analysis uses values for 4th-grade students at urban, low-socioeconomic schools in the western United States. We consulted Schochet (2008) for the estimated range of ICCs at the teacher level (ρ_2). Table 7 shows the minimum detectable effect size (MDES) at 80% power with an alpha threshold of 0.05 for the various sets of assumptions.

Table 7. Minimum detectable effect sizes

R^2	ρ_3	ρ_2	**MDES**
Math standardized tests			
0.876	0.195	0.1	0.35
		0.2	0.38

ELA standardized tests			
0.607	0.157	0.1	0.38
		0.2	0.45

We do not expect to find effect sizes this large on the distal outcome measures.According to Lipsey et al. (2012), these effects would be roughly equivalent to two trimesters of growth for math and a year or more of growth for ELA. However, we believe effect sizes in the range of 0.35-0.45 (moderate effects) are plausible for the LCTAA and ICSM because they are non-standardized instruments, where effect sizes are typically much larger (Cheung & Slavin, 2015). We believe the study is sufficiently powered to detect effects on the proximal measures, assuming that R^2, ρ_3, and ρ_2 for the LCTAA and ICSM are similar to the values above.

4. Key project components, mediators, outcomes, and measurable threshold

Outcomes. To estimate the effectiveness of the intervention on the various outcomes, we will use the following three-level hierarchical linear model (HLM), with students at level 1 (i), teachers at level 2 (j), and school at level 3 (k):

Level 1: $Outcome_{ijk} = \beta_{0jk} + \beta_{1jk}Pre_{ijk} + \Sigma\beta_1 Demo_{ijk} + \varepsilon_{ijk}$

Level 2: $\beta_{0jk} = \gamma_{00k} + \Sigma\gamma_J Experience_{jk} + \eta_{0jk}$

Level 3: $\gamma_{00k} = \delta_{000} + \delta_{001}Tx_k + \delta_{002}Pre_k + \Sigma\delta_K Demo_k + \xi_{00k}$

In this model, *Outcome* represents the post-test measure of interest. At the student level, *Pre* represents the student's pre-test score on the measure (or, in the case of the SBAC, the student's scores from the end of 3rd grade). At the school level, *Pre* represents the school's average pre-test score. *Demo* represents a vector of student-level (i.e., gender, ethnicity, English fluency, and participation in special education programs) or school-level demographic characteristics (i.e., ethnic makeup of the student body, proportion of low-income students, and proportion of students who are ELLs). *Experience* represents a vector of teacher-level variables that includes educational background and prior teaching experience. ε_{ijk} represents student-level error. η_{0jk} and

ξ_{00k} represent random effects of teacher and school, respectively. In this model, the main effect of the intervention is captured by δ_{001}.

Implementation data. Quarterly classroom observations will continue during Year 4. In Year 5, we will observe 8 teachers. When teachers implement the curriculum, they will create online accounts in Scratch for their classes, where students will post their projects in a separate "studio" for each module of the curriculum. Researchers will be given access to the online class accounts. Researchers will determine, at least quarterly, how many units of curriculum have been taught based on the number of classroom studios with projects.

In addition, researchers will measure specific and generic fidelity through regular electronic surveys. Specific fidelity will be measured with a weekly survey of treatment teachers to determine whether they taught the treatment curriculum that week and, if so, for how many minutes and which module. Generic fidelity will be measured with a monthly survey of treatment and control or delayed-treatment teachers to determine if they taught any CS curriculum (other than the treatment curriculum) during the prior month, and if so for how many minutes. This will provide data not only on implementation, but on the achieved relative strength (Hulleman & Cordray, 2009) of the curriculum and enable analyses to explore number of units taught and minutes spent teaching the treatment curriculum as potential mediators of our outcomes in the RCT.

Based on the results of the impact analysis, WestEd will generate 4 strata of teachers in the treatment condition based on prior performance and the amount of learning gain of their students on multiple outcome measures. WestEd researchers will examine the homogeneity of each subgroup and analyze the characteristics of each stratum in terms of the frequency and amount of use of the intervention, use of different modules in the intervention, demographics of the students, students' prior performance, etc. Based on the analysis, we will strategically sample 2 teachers from each stratum to observe their classrooms and follow up with an interview in Year 5 to take a closer look at the implementation characteristics and settings that led to more or less student learning in order to help improve the intervention and its implementation quality, and inform replicability and scaling.

Moderators. Moderator analysis will be performed by extending the above model, representing moderator effects as interactions between the treatment and the moderator variable of interest (Baron & Kenny, 1986). To estimate moderator effects between treatment and school-level variables, we will add an interaction term to Level 3. To estimate moderator effects between treatment and student-level or teacher-level variables, we will add additional equations to Level 2 and/or Level 3 of the model to predict the coefficient of the variable of interest.

Mediators. Analysis of fidelity as a mediator will be performed using structural equation modeling. We will first test the effect of treatment on the amount of CS curriculum (treatment and control) students experienced. Then we will test the effect of CS curriculum on student outcomes, controlling for treatment. Finally, we will test the effect of treatment directly on student outcomes. The significance of the mediation effect will be determined using the Empirical M-test and PRODCLIN software (MacKinnon, Fritz, Williams, & Lockwood, 2007). Examining fidelity as a mediator of student outcomes will yield useful information regardless of a significant overall effect of treatment. If treatment fails to predict amount of CS instruction, then the lack of effect was due to lack of program differentiation, suggesting that either control teachers regularly incorporate computational thinking activities into their typical practice or that treatment teachers needed additional support to implement the treatment with fidelity.

References

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182. doi:10.1037/0022-3514.51.6.1173

Barron, B., Martin, C. K., Takeuchi, L., & Fithian, R. (2009). Parents as learning partners in the development of technological fluency. *International Journal of Learning and Media, 1*(2), 55-77. doi:10.1162/ijlm.2009.0021

Bennedsen, J., & Caspersen, M. E. (2012). Persistence of elementary programming skills. *Computer Science Education*, *22*(2), 81-107. doi:10.1080/08993408.2012.692911

Brown, B. A., & Ryoo, K. (2008). Teaching science as a language: A “content-first” approach to science teaching. *Journal of Research in Science Teaching*, *45*(5), 529-553.

Brown, J. E., & Doolittle, J. (2008). A cultural, linguistic, and ecological framework for response to intervention with English language learners. *Teaching Exceptional Children*, *40*(5), 66-72.

Bureau of Labor Statistics. (2018). *Occupational outlook handbook: Computer and information technology*. Retrieved from https://www.bls.gov/ooh/computer-and-information-technology/home.htm

Burke, Q., & Kafai, Y. B. (2012, February). The writers' workshop for youth programmers: Digital storytelling with Scratch in middle school classrooms. *Proceedings of the 43rd ACM Technical Symposium on Computer Science Education*, 433-438. ACM.

Carlone, H. B., Scott, C. M., & Lowder, C. (2014). Becoming (less) scientific: A longitudinal study of students' identity work from elementary to middle school science. *Journal of Research in Science Teaching*, *51(7)*, 836-869. doi:10.1002/tea.21150

Cheung, A. C. K., & Slavin, R. E. (2015). *How methodological features affect effect sizes in education.* Baltimore, MD: Johns Hopkins University, Center for Research and Reform in Education.

Childs, R. A., & Jaciw, A. P. (2003). Matrix sampling of items in large-scale assessments. *Practical Assessment, Research & Evaluation*, *8*(16), 1–9.

College Board. (2017). AP program participation and performance data 2017 [Data file]. Retrieved from https://research.collegeboard.org/programs/ap/data/participation/ap-2017

de Ayala, R. (2010). Item response theory. In G. R. Hancock, R. O. Mueller, & L. M. Stapleton (Eds.), *The Reviewer's Guide to Quantitative Methods in the Social Sciences*, 155. New York, NY: Routledge.

de Souza, C., Garcia, A., Slaviero, C., Pinto, H., & Repenning, A. (2011). Semiotic traces of computational thinking acquisition. *End-User Development*, 155-170.

Donker, A. S., de Boer, H., Kostons, D., Van Ewijk, C. D., & van der Werf, M. P. (2014). Effectiveness of learning strategy instruction on academic performance: A meta-analysis. *Educational Research Review*, *11*, 1-26. doi:10.1016/j.edurev.2013.11.002

Estrella, G., Au, J., Jaeggi, S. M., & Collins, P. (2018). Is inquiry science instruction effective for English language learners? A meta-analytic review. *AERA Open, 4*(2), 1-23. doi:10.1177/2332858418767402

Fancsali, C., Tigani, L., Toro Isaza, P., & Cole, R. (2018, February). A landscape study of computer science education in NYC: Early findings and implications for policy and practice. *SIGCSE '18 Proceedings of 49th ACM Technical Symposium on Computer Science Education,* 44-49. New York, NY: ACM. doi:10.1145/3159450.3159467

Gilmartin, S. K., Li, E., & Aschbacher, P. (2006). The relationship between secondary students' interest in physical science or engineering, science class experiences, and family contexts: Variations by gender and race/ethnicity. *Journal of Women and Minorities in Science and Engineering, 12*(2–3), 179–207.

Gomez-Zwiep, S. (2017, March). *Creating equitable STEM spaces for English learners*. Paper presented at the California STEM Includes Conference, Anaheim, CA.

Goode, J., & Margolis, J. (2011). Exploring computer science: A case study of school reform. *ACM Transactions on Computing Education (TOCE)*, *11*(2), 12. doi:10.1145/1993069.1993076

Grunow, A. (2015). Improvement discipline in practice [Blog Post]. Retrieved from https://www.carnegiefoundation.org/blog/improvement-discipline-in-practice/

Hedges, L. V., & Hedberg, E. C. (2007). Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis, 29*(1), 60-87. doi:10.3102/0162373707299706

Hulleman, C. S., & Cordray, D. S. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness, 2*(1), 88-110.

Israel, M., Pearson, J. N., Tapia, T., Wherfel, Q. M., & Reese, G. (2015). Supporting all learners in school-wide computational thinking: A cross-case qualitative analysis. *Computers & Education*, *82*, 263-279.

Jacob, S., Garcia, L., & Warschauer, M. (2018, December). *Leveraging multilingual identities in computer science education.* Paper presented to the California Teaching English to Speakers of Other Languages Los Angeles Regional Conference, Los Angeles, CA.

Jacob, S., Nguyen, H., Garcia, L. Richardson, R., & Warschauer, M. (2019, February).

Developing a computational thinking curriculum for multilingual students. Poster presented at RESPECT 2019, IEEE Special Technical Community on Broadening Participation, Minneapolis, MN.

Jacob, S., Nguyen, H., Tofel-Grehl, C., Richardson, D., & Warschauer, M. (2018). Teaching computational thinking to English learners. *NYS TESOL Journal*, *5*(2), 12-24.

Jacob, S. R., & Warschauer, M. (2018). Computational thinking and literacy. *Journal of Computer Science Integration*, *1*(1), 1-19.

Janzen, J. (2008). Teaching English language learners in the content areas. *Review of Educational Research, 78*(4), 1010-1038.

Johnson, L., Graham, S., & Harris, K. R. (1997). The effects of goal setting and self-instruction on learning a reading comprehension strategy: A study of students with learning disabilities. *Journal of Learning Disabilities, 30*(1), 80-91.

Kelleher, C., & Pausch, R. (2007). Using storytelling to motivate programming. *Communications of the ACM, 50*(7), 58-64.

Kelly, G. J., & Chen, C. (1999). The sound of music: Constructing science as sociocultural practices through oral and written discourse. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *36*(8), 883-915.

Kena, G., Musu-Gillette, L., Robinson, J., Wang, X., Rathbun, A., Zhang, J., ... & Velez, E. D. V. (2015). The condition of education 2015 (NCES 2015-144). Retrieved from National Center for Education Statistics website: https://nces.ed.gov/pubs2015/2015144.pdf

Lee, I., Martin, F., Denner, J., Coulter, B., Allan, W., Erickson, J., ... & Werner, L. (2011). Computational thinking for youth in practice. *ACM Inroads*, *2*(1), 32-37.

doi:10.1145/1929887.1929902

Lee, O., Llosa, L., Grapin, S., Haas, A., & Goggins, M. (2019). Science and language integration with English learners: A conceptual framework guiding instructional materials development. *Science Education, 103*(2), 317-337. doi: 10.1002/sce.21498

Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., ... & Busick, M. D. (2012). Translating the Statistical Representation of the Effects of Education Interventions into More Readily Interpretable Forms (NCSER 2013-3000). *National Center for Special Education Research*.

MacKinnon, D. P., Fritz, M. S., Williams, J., & Lockwood, C. M. (2007). Distribution of the product confidence limits for the indirect effect: Program PRODCLIN. *Behavioral Research Methods, 39*(3), 384.

Manz, S. L. (2002). A strategy for previewing textbooks: Teaching readers to become THIEVES (Teaching ideas). *The Reading Teacher*, *55*(5), 434-436.

Martin, A., McAlear, F., & Scott, A. (2015). Path not found: Disparities in access to computer science courses in California high schools. Retrieved from ERIC: https://eric.ed.gov/?id=ED561181

McGinn, M. K., & Roth, W. M. (1999). Preparing students for competent scientific practice: Implications of recent research in science and technology studies. *Educational Researcher*, *28*(3), 14-24.

Michaels, S., & O’Connor, C. (2015). Conceptualizing talk moves as tools: Professional development approaches for academically productive discussions. In L. B. Resnick, C. Asterhan, & S. Clarke (Eds.), *Socializing intelligence through academic talk and dialogue* (pp. 347-362).

National Academies of Sciences, Engineering, and Medicine. (2018). *English learners in STEM subjects: Transforming classrooms, schools, and lives*. Washington, DC: The National Academies Press. doi:10.17226/25182

National Research Council. (1996). National science education standards. *ERIC Document Reproduction Service, 391*(690). Washington, DC: National Academy Press.

National Research Council. (2012). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.

National Research Council. (2014). *STEM integration in K-12 education: Status, prospects, and an agenda for research*. Washington, DC: The National Academies Press. doi:10.17226/18612

NWEA Psychometrics Service Team. (2017). *Linking the Smarter Balanced Assessments to NWEA MAP Growth Tests*. Retrieved from https://www.nwea.org/resources/linking-the-smarter-balanced-assessments-to-nwea-map-assessments/

Palincsar, A. S. (1986). Metacognitive strategy instruction. *Exceptional Children*, *53*(2), 118-124.

Peppler, K. A., & Kafai, Y. B. (2007). From SuperGoo to Scratch: Exploring creative digital media production in informal learning. *Learning, Media and Technology, 32*(2), 149-166.

Peppler, K., & Warschauer, M. (2012). Uncovering literacies, disrupting stereotypes: Examining the (dis)abilities of a child learning to computer program and read. *International Journal of Learning and Media*, *3*(3), 15-41.

Pressley, M. (2002). Metacognition and self-regulated comprehension. In A. E. Farstrup & S. J. Samuel (Eds.), *What research has to say about reading instruction* (3rd ed., pp. 291-309). doi:10.1598/0872071774.13

Raudenbush, S. W., Spybrook, J., Congdon, R., Liu, X., Martinez, A., Bloom, H., & Hill, C. (2011). Optimal Design Plus Empirical Evidence (Version 3.0) [Software]. Retrieved from www.wtgrantfoundation.org

Rosebery, A. S., & Warren, B. (Eds.). (2008). *Teaching science to English language learners: Building on students' strengths.* Arlington, VA: National Science Teachers Association.

Royal, D., & Swift, A. (2016, October 18). U.S. minority students less exposed to computer science. *Gallup.* Retrieved from http://www.gallup.com/poll/196307/

Ryoo, K., Bedell, K., & Swearingen, A. (2018). Promoting linguistically diverse students' short-term and long-term understanding of chemical phenomena using visualizations. *Journal of Science Education and Technology*, *27*(6), 508-522.

Sáez-López, J. M., Román-González, M., & Vázquez-Cano, E. (2016). Visual programming languages integrated across the curriculum in elementary school: A two year case study using "Scratch" in five schools. *Computers & Education, 97*, 129-141.

Salac, J., White, M., Wang, A., & Franklin, D. (2019, February). An analysis through an equity lens of the implementation of computer science in K-8 classrooms in a large, urban school district. *Proceedings of the 50th ACM Technical Symposium on Computer Science Education,* 1150-1156. ACM.

Schochet, P. Z. (2008). Statistical power for random assignment evaluations of education programs. *Journal of Educational and Behavioral Statistics, 33*(1), 62-87. doi:10.3102/1076998607302714.

Schunk, D. H. (2003). Self-efficacy for reading and writing: Influence of modeling, goal setting, and self-evaluation. *Reading & Writing Quarterly*, *19*(2), 159-172.

Shea, L. M., & Shanahan, T. B. (2011). Methods and strategies: Talk strategies. *Science and*

Children, *49*(3), 62-66.

Shea, L. M., Sandholtz, J. H., & Shanahan, T. B. (2017). We are all talking: A whole-school approach to professional development for teachers of English learners. *Professional Development in Education*, *44*(2), 190-208.

Shettle, C., Roey, S., Mordica, J., Perkins, R., Nord, C., Teodorovic, J., ... & Brown, J. (2007). The nation's report card: America's high school graduates. *National Center for Education Statistics*, *467*.

Smarter Balanced Assessment Consortium. (n.d.). 2016-2017 technical report. Retrieved from http://www.smarterbalanced.org/assessments/development/

von Davier, A. A., & Wilson, C. (2007). IRT true-score test equating: A guide through assumptions and applications. *Educational and Psychological Measurement, 67*(6), 940–957. doi:10.1177/0013164407301543

Wang, J., Hong, H., Ravitz, J., & Hejazi Moghadam, S. (2016, February). Landscape of K-12 computer science education in the US: Perceptions, access, and barriers. *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, 645-650. Memphis, TN: ACM. doi:10.1145/2839509.2844628

Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., & Wilensky, U. (2016). Defining computational thinking for mathematics and science classrooms. *Journal of Science Education and Technology, 25*(1), 127-147.

Werner, L., Denner, J., Campe, S., & Kawamoto, D. C. (2012, February). The fairy performance assessment: Measuring computational thinking in middle school. *Proceedings of the 43rd ACM Technical Symposium on Computer Science Education*, 215-220. Raleigh, NC: ACM. doi:10.1145/2157136.2157200

Wiebe, E., London, J., Aksit, O., Mott, B. W., Boyer, K. E., & Lester, J. C. (2019, February). Development of a lean computational thinking abilities assessment for middle grades students. *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, 456-461. ACM.

Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, *49*(3), 33-35.

Wright, B. D. (1977). Solving measurement problems with the Rasch model. *Journal of Educational Measurement*, *14*(2), 97–116. doi:10.1111/j.1745-3984.1977.tb00031.x

Zwiep, S. G., & Straits, W. J. (2013). Inquiry science: The gateway to English language proficiency. *Journal of Science Teacher Education, 24*(8), 1315-1331.