Validating the SunBay Middle School Digital Mathematics Program:
An Innovative Approach to Increasing Student Achievement Using Technology

Contents

A. Significance.......................................................................................................................... 2
  A.1 Likelihood of Estimated Impact .................................................................................... 3
      a. Strong theoretical foundation....................................................................................... 3
      b. Moderate evidence of effectiveness from prior research............................................. 3
      c. Evidence of positive impact for high-need students.................................................... 4
      d. Evidence of replicability of the SunBay Mathematics approach.................................. 4
      e. Strong consonance with the PARCC assessment.......................................................... 4
  A.2 Feasibility of National Expansion ................................................................................. 5
B. Quality of the Project Design ............................................................................................... 6
  B.1 Addressing a National Need via the SunBay Approach ................................................ 6
  B.2 The Design of SunBay Mathematics for the Common Core and PARCC .................... 7
  B.3 Our Logic Model and Strategy for Addressing Barriers ............................................. 10
  B.4 This Project’s Goals and Plan of Action ...................................................................... 11
  B.5 Competitive Priority 1: Improving Cost-Effectiveness and Productivity .................... 15
  B.6 Competitive Priority 2: Enabling Broad Adoption of Effective Practices .................. 17
C. Management Plan ............................................................................................................... 18
  C.1 Roles and Responsibilities ........................................................................................... 19
  C.2 Timelines and Milestones ............................................................................................ 20
  C.3 Metrics and Performance Targets ................................................................................ 22
  C.3 Financial and Operating Model: Plan to Operate at a Regional Level ........................ 22
D. Personnel ............................................................................................................................ 23
  D.1. Coverage of All Key Positions ................................................................................... 23
  D.2. Qualifications and Expertise of Project Director and Key Personnel ........................ 24
E. Project Evaluation .............................................................................................................. 24
  E.1 Key Questions Addressed in the Project Evaluation and Methods Used .................... 25
  E.2 How Evidence Produced Will Meet WWC Evidence Standards................................... 25
  E.3 Evaluating at the Proposed Level of Scale: Sample and Random Assignment .......... 25
  E.4 Measures, Power Analysis, and Analysis Plan ............................................................ 27
      a. Measures. .................................................................................................................... 27
      b. Power analysis. ........................................................................................................... 30
      c. Analysis plan. .............................................................................................................. 31
  E.5. Outcomes of Project and Measures of Implementation .............................................. 35
References............................................................................................................................... 36
Validating the SunBay Middle School Digital Mathematics Program:
An Innovative Approach to Increasing Student Achievement Using Technology

SRI International will lead a validation project to address a critical need for middle school students and teachers: increased achievement on the Common Core State Standards for Mathematics (CCSSM) as measured by the Partnership for Assessment of Readiness for College and Career (PARCC) assessments. This project will evaluate SunBay Mathematics, an innovative technology-based approach to middle school mathematics that has been developed through an extensive program of prior research. The design of SunBay Mathematics is deeply grounded in Learning Science principles, and the research-based development process accurately anticipated the CCSSM focus on conceptual understanding and mathematical practices, as well as CCSSM’s key middle school mathematics topics. Moderate evidence of effectiveness for SunBay Mathematics was established in an NSF-funded evaluation, and positive impacts for high-need students were pronounced on the measures available at the time. Findings were replicated in schools in Texas, Florida, and England, showing potential for national impact.

This project will validate SunBay Mathematics as an integrated technology-rich approach to the implementation of CCSSM while measuring both student outcomes (PARCC) and teacher outcomes [Absolute Priority 4 in area (c)]. The project seeks to improve outcomes while reducing costs (Competitive Priority 1) by increasing collaborative learning on shared tablet devices, shifting from paper to digital workflows, and blending online and in-person teacher professional development. The project also enables broad adoption by strengthening SunBay’s modular approach to introducing effective teaching practices (Competitive Priority 2).

This project addresses the key barriers to scaling SunBay Mathematics. To overcome the evidentiary barrier, we will provide the validity evidence needed for districts, schools, and teachers using the PARCC measures, which are more comprehensive than the assessments in our prior research. We address the cost barrier by developing and studying lower cost models of implementation. We address the adoption barrier by streamlining the use of 2- to 3-week-long modules to introduce effective teaching practices with technology. We address these barriers
through partnerships with school districts in Florida serving high-need students (Broward and Palm Beach counties), university-based teacher professional development programs [Lastinger Center at the University of Florida (UF) and the University of South Florida St. Petersburg (USFSP)], and an external evaluator (the Consortium for Policy Research in Education).

**A. Significance**

**Important need.** Middle school mathematics is important to students’ future success as it is the gateway to college readiness and to coursework needed for science, technology, engineering, and mathematics careers (Evan, Gray, & Olchefske, 2006). Florida and 44 other states have adopted the CCSSM, which concentrate on deep understanding of fewer key ideas each year than current standards and emphasize students’ conceptual understanding and skill in mathematical practices. Most existing classroom materials and approaches do not match the focus areas of the CCSSM because they were previously not measured by statewide assessments.

**Unmet demand creates opportunity for scaling up.** Significant unmet demand exists using technology in integration with curriculum materials and teacher professional development to implement the CCSSM. Specifically, Florida, the District of Columbia, and 21 other states plan to use PARCC assessments as they implement CCSSM. Considering only those students served by both CCSSM standards and PARCC assessments, approximately 3 million middle school students per year could benefit from SunBay Mathematics. Preparing for the CCSSM and PARCC is a top concern of the school districts serving these students, partly because PARCC assessments will use technology-based items that are not well aligned with traditional curriculum materials. For example, a recent American Federation of Teachers poll\(^1\) showed that 74% of teachers are worried that CCSSM-aligned assessments will precede availability of materials and teacher readiness to teach to the standards. Further, PARCC assessments will be delivered via digital technology, so students must be technologically proficient. Consequently, demand for such materials is likely to be high in Florida (our immediate target) and nationally.

In choosing new materials to match these assessments, school districts want evidence of effectiveness, and research on technology-rich curriculum materials has not always found strong evidence of effectiveness despite considerable experimentation with materials that were presumed to be the best available (Cheung & Slavin, 2011; Dynarski et al., 2007). However, SunBay Mathematics has demonstrated consistently positive results. This project will extend that evidence to the new CCSSM-aligned PARCC assessments. From prior research, we estimate SunBay’s impact on student achievement to be at least .20 standard deviations; this is the equivalent of moving a student from the 50th to the 58th percentile; sustained over 3 years, this would be equivalent to moving a student from the 50th to the 73rd percentile.

A.1 Likelihood of Estimated Impact

SunBay Mathematics will produce its estimated impact on student achievement because of its (a) strong theoretical foundation, (b) moderate evidence of effectiveness from prior research, (c) evidence of impacts for high-need students, (d) replicable modular adoption model, and (e) consonance with the PARCC assessment.

a. Strong theoretical foundation. Both SunBay and the CCSSM focus on key mathematical ideas and important mathematical practices, consistent with recommendations by the National Council of Teachers of Mathematics (2006) and the National Mathematics Advisory Panel (2008) and the standards in CCSSM. In middle school, the CCSSM key ideas and mathematical practices center on the transition from arithmetic to algebra—on ratios, proportional reasoning, and expressions and equations. SunBay integrates research-based technology in curricular materials to address these topics, aligning strong theories for curriculum design, technology integration, and teacher professional development. See Section B.2 for detail on the design principles that establish SunBay’s program theory.

b. Moderate evidence of effectiveness from prior research. Moderate evidence for effectiveness of SunBay Mathematics is provided by experiments in Texas, which met the What Works Clearinghouse (WWC) evidence standards without reservations (SRI renamed the
program SunBay from SimCalc in 2009; see http://tinyurl.com/sunbaywwc). Per WWC Evidence Standards (Department of Education, 2008), the two experiments each used a randomized controlled trial (RCT) design. Attrition was low (19%) and not different between groups (Tatar & Stroter, 2009). Groups were equivalent at pretest. The contrast in each experiment avoided confounding factors (Roschelle et al., 2010). The validity and reliability of the outcome measures were established (Shechtman, Haertel, Roschelle, et al., 2010). A peer-reviewed article (Roschelle et al., 2010) addressed the WWC caution about potential overalignment of our measures, and we have since shown evidence of their external validity by examining correlations in student performance on a state-administered measure (see Appendix D).

c. Evidence of positive impact for high-need students. The overall student-level effect sizes reported for the Texas experiments were .63, .50, and .56, using end-of-unit proximal measures. SunBay was designed to support high-need students, and the analyses (Roschelle et al., 2010) found that positive impacts were robust across the settings and populations studied, which included students in poverty, students at high-minority schools, and students with low prior achievement (see Appendix D).

d. Evidence of replicability of the SunBay Mathematics approach. After the Texas experiment, we took SunBay Mathematics to Florida with the opening of SRI’s St. Petersburg office, developed teacher professional development (TPD) capacity in the state, and conducted quasi-experiments with the support of the Helios Education Foundation. (Texas is not adopting the CCSSM, so conducting a validation study there would limit generalizability.) Analyses in Florida showed the same learning gains as in Texas and revealed no detectable differences between the Florida and Texas treatment groups’ pretest and gain scores. In addition, we found significant learning gains for African American students, a group underrepresented in our Texas study. A follow-up analysis also showed evidence of a clear trend linking gains on SunBay unit tests to gains on Florida’s state test, the FCAT (see Appendix D).

e. Strong consonance with the PARCC assessment. The PARCC online assessments “require students to express their mathematical reasoning and to apply key mathematical skills,
concepts, and processes to solve complex problems.” Exhibit 1 shows a sample item requiring comparison of speeds across graphs and tables. Overall, alignment with released items is strong (but as only a sample of items have been released, a full alignment study is not yet possible). Whereas traditional textbook materials do not prepare students for this type of item, SunBay builds on students’ direct perceptual experience of motion in technology-based simulations, providing experiences in comparing speeds across graphs and tables, as shown in Exhibit 2.

Exhibit 1: PARCC sample (http://www.parcconline.org/samples/mathematics/grade-7-Speed)

Exhibit 2: SunBay computer activity and workbook page comparing speed in graphs and tables

A.2 Feasibility of National Expansion

We will use SRI’s Discipline of Innovation (Carlson & Wilmot, 2007) process to bring SunBay Mathematics to scale. Through this process, SRI has developed innovations that began as basic research, proceeded to pilots and prototypes, and then moved into the national market. A well-
known current example is Siri, the iPhone virtual personal assistant. Examples of SRI’s capacity to achieve national scope in education are our EduSpeak technology (used to teach English as a second language) and incubation of Teachscape.com (an online TPD provider).

Institutionally, with support from the Helios Education Foundation, this project will accelerate the team’s readiness to scale nationally by establishing SRI’s St. Petersburg Center for Digital Learning as the operations center for SunBay scale-up. This center, seeded with $1 million from the foundation, was created to build a portfolio of digital learning projects in Florida to enhance student achievement in STEM. Our national expansion model combines SRI’s proven ability to scale technology to a national scale with university partnerships to provide professional development services in each region. For example, in Florida our primary PD partner, the Lastinger Center, serves all of Florida with its Algebra Nation program, an innovative approach to supporting teachers and students through video, social media, and community content. With i3 funding, we will fully shift our technology to a scalable cloud infrastructure and strengthen, test, and refine joint operating procedures with university-based TPD providers. In particular, our approaches to Competitive Priorities 1 and 2, described later, are firmly aligned with the goal of program expansion. Once favorable outcomes are found in this validation study, we will be ready for further expansion in Florida and the nation.

B. Quality of the Project Design

B.1 Addressing a National Need via the SunBay Approach

SunBay Mathematics addresses the critical national need to increase learning for all students in rigorous middle school mathematics, considered a gateway to college readiness and to STEM careers (Evan et al., 2006). Schools need resources to tackle the heightened focus on conceptual understanding and mathematical practices in the CCSSM and to thereby increase student achievement on PARCC technology-based assessments. SunBay effectively addresses the full extent of this challenge via its systems integration approach to effective learning. As described in Section B.2, we integrate technology, curriculum, and professional development. SunBay
Mathematics has an established capability to enable 90% of teachers to achieve high-quality implementation across diverse settings and with a wide range of students (see Appendix D).

For implementation with diverse teachers in varied settings, schools need systems that are easy to adopt, that diverse teachers can implement with high fidelity, and that can be adapted and expanded to fit site-specific needs. SunBay’s modular approach meets these needs as it helps teachers address their most difficult teaching challenges. Each 2- to 3-week module emphasizes a key mathematical idea, has associated ongoing PD, includes clear adaptation guidelines and processes, and gives teachers data to continually refine their practices and the program. SunBay builds commitment within schools by demonstrating results and building teacher leaders’ capacity to assume ownership of the program; once they have mastered the basic modules, these leaders can adapt and expand SunBay. In addition to strong interest for expansion throughout Florida (see letter from FL State Representative Brandes and local industry leader Betzer), schools in Virginia and California have also asked to use SunBay mathematics. The support for adoption, implementation and adaptation in SunBay Mathematics strongly addresses what schools are looking for as they implement CCSSM.

B.2 The Design of SunBay Mathematics for the Common Core and PARCC

The SunBay technology design emphasizes linked dynamic representations such as graphs that are linked to motions, tables, and equations (Kaput, Hegedus, & Lesh, 2007; Vahey, Roy, & Fueyo, 2013; see Exhibit 2). While many evaluations show at best small positive effects for technology (Cheung & Slavin, 2011; Dynarski et al., 2007), dynamic representations have been shown to be powerful in developing conceptual understanding (Heid & Blume, 2008), and consistent with the Multimedia Principle (Fletcher & Tobias, 2005), they help students make important connections between intuitive and formal ideas and between graphical and linguistic understandings. Increasing students’ conceptual understanding and making such connections are fundamental to the CCSSM. Further, dynamic representations provide exceptional opportunities for students to engage in mathematical practices, and our research has demonstrated
effectiveness in increasing those practices in the classroom (Shechtman, Knudsen, & Stevens, 2010). *SunBay's technology design emphasizes technology uses, such as linked dynamic representations, that have proven potential to increase student learning in exactly the ways necessary for increased achievement on the CCSSM-aligned and PARCC assessments.*

The *SunBay curriculum module design* is based on organizing key mathematical ideas and important mathematical practices in the CCSSM (see Exhibit 3) into materials that are easy to use in classrooms. These modules integrate core ideas from Singapore’s mathematics education research that resonate with CCSSM, including the balance of concepts and procedures, careful choice of representations, and support for all students to solve challenging problems (Yoong et al., 2009). To ensure that our materials are effective for high-need students, we employ principles from Universal Design for Learning (Rose, Meyer & Hitchcock, 2005), such as

<table>
<thead>
<tr>
<th>CCSSM Content</th>
<th>Focus Practices</th>
<th>SunBay Theme</th>
<th>Representations</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.RP: Understand ratio concepts and use ratio reasoning to solve problems.</td>
<td>Reason abstractly and quantitatively.</td>
<td>Music and video: e.g., rates in beats per minute; % time elapsed vs. whole in video</td>
<td>Simulations, tables, double number lines</td>
</tr>
<tr>
<td>6.EE …extend understandings of arithmetic to algebraic expressions</td>
<td>Argumentation; Look for and make use of structure and regularity.</td>
<td>Design with patterns: e.g., establish addition of like terms as way to show equivalence of expressions</td>
<td>Pattern generator; tables; expression manipulation</td>
</tr>
<tr>
<td>7.RP: Analyze proportional relationships and use them to solve real-world and mathematical problems.</td>
<td>Argumentation; Model with mathematics.</td>
<td>Managing a Soccer Team: e.g., calculating rates of running, buying equipment</td>
<td>Motion simulations, graphs, tables, equations</td>
</tr>
<tr>
<td>7.EE: Use properties of operations to generate equivalent expressions</td>
<td>Argumentation; Look for and make use of structure and regularity.</td>
<td>Design with patterns: e.g., establish factoring as way to show equivalence of expressions</td>
<td>Pattern generator, expressions linked to pattern components</td>
</tr>
<tr>
<td>8.EE: Connections between proportional relationships, lines, and linear equations.</td>
<td>Argumentation; Model with mathematics.</td>
<td>Design mobile games: e.g., compare speeds, including nonlinear and “head starts”; map sales to profits</td>
<td>Motion simulations, graphs, tables, equations (see Exhibit 2)</td>
</tr>
<tr>
<td>8.G: Understand congruence and similarity using physical models or geometry software.</td>
<td>Argumentation; Use appropriate tools.</td>
<td>Design an online zine: e.g., fit images, geometric objects, etc., on different screens without distortion</td>
<td>Dynamic geometry, including dilating, rotating, and reflection</td>
</tr>
</tbody>
</table>
emphasizing meaningful representations, supporting students’ action, and inclusive teaching strategies. Within the modules, students solve increasingly challenging problems in one content area organized by a high-interest theme such as “Designing Mobile Games.” Each module covers a small number of practices, and during each class students engage in a combination of whole-class, small-group, and individual learning activities. Teachers focus attention on the key mathematical ideas, emphasizing conceptual understanding and guiding development of mathematical practices. The module design aligns with our technology design and with the CCSSM, providing curricular learning progressions and teacher-guided classroom activities designed to enable all students to meet rigorous expectations.

The SunBay TPD design emphasizes practical “mathematical knowledge for teaching” (Hill et al., 2008), which is the knowledge that teachers need to make sense of and extend their students’ mathematical reasoning. We also emphasize Technological Pedagogical Content Knowledge (Mishra & Koehler, 2006), which is the knowledge teachers need to teach effectively with technology. In each module, we emphasize a small set of teaching practices (Ball & Forzani, 2009; Hammerness, 2006), with a focus on supporting argumentation in the classroom (Kim, 2012). This is consistent with the CCSSM recognition that the ability to “construct viable arguments and critique the reasoning of others” is a core practice and that “one hallmark of mathematical understanding is the ability to justify.” Our TPD approach has measurable impacts in students’ participation in extended argumentation (Shechtman, Knudsen, & Stevens, 2010).

Our TPD design also incorporates research-based principles of long-term, ongoing TPD (Banilower et al., 2006), forming a professional learning community (PLC) (Wei et al., 2009), and engaging teachers in reflective practice with new teaching materials, technologies, and moves (Darling-Hammond, 2008). To take full advantage of our PLCs, besides providing TPD to introduce all teachers to the CCSSM and SunBay materials, we provide certificates or master’s degrees to “teacher leaders” as they show advanced knowledge of the materials, which prepares them to lead local PLCs. Thus we address the absolute priority requirement of teaching efficacy
by systematically building capacity across a district’s teaching workforce to use technology-based materials to implement the CCSSM and increase student achievement on PARCC.

B.3 Our Logic Model and Strategy for Addressing Barriers

The logic model (Exhibit 4) guided the design of this validation project. All the work will be conducted in the context of college- and career-ready standards as described in CCSSM and measured by PARCC. Work will occur in two settings in Florida. In both settings work will flow from design to well-defined inputs that are then implemented in classrooms to achieve outcomes.

Exhibit 4: Logic model

The first setting will include two counties (Palm Beach and Broward), where we will conduct an RCT pursuant to Absolute Priority 4. The design work will include refinements of the modules to meet district requirements and modifications of district documents and policies (such as pacing guides and TPD schedules) to guide teachers in implementing SunBay while meeting district requirements. This work will result in three kinds of inputs for the SunBay modules: content (curriculum materials), technology (software used in the curriculum materials), and TPD (both for all teachers and for teacher leaders). The resulting modules will then be used in
classrooms as an implementation of the CCSSM. Outcomes will be measured for students in grades 6, 7, and 8 and for teachers per the plan in the Evaluation section. As lessons are learned during implementation and initial analyses, insights will be fed back into the design process.

The second setting will host our additional studies for Competitive Priorities 1 and 2. These will occur in additional Florida districts and schools recruited for this purpose (called Competitive Priorities schools, or CP schools). We will vary the design of each input, seeking improvements in cost-effectiveness, productivity, and scalability. We will measure student and teacher outcomes as well as cost, productivity, and scalability outcomes. At the end of the work in both settings, all that we have learned will be synthesized in preparation for further scaling.

This logic model represents our strategy for addressing three barriers:

1. **Evidentiary barrier**—by producing the evidence for educational decision-makers that SunBay Mathematics increases student achievement on measures of the college- and career-readiness, using the PARCC assessments as the primary outcome measure in an RCT.

2. **Cost barrier**—by developing and studying lower cost implementation models. Design work will vary each input in our logic model in an effort to reduce costs and increase productivity. Results will be measured and iteratively improved.

3. **Adoption barrier**—by readying the practice of using modules to introduce effective teaching practices with technology.

**B.4 This Project’s Goals and Plan of Action**

Our project has five goals. Goals 2 and 3 support Competitive Priorities. Timelines, milestones, and responsibilities for all goals and actions are in the Management and Evaluation plans.

**Goal 1: Increase student achievement in Florida schools by integrating the SunBay technology-based materials into district efforts to implement the rigorous CCSSM standards.** This goal is necessary both for Absolute Priority 4 and the Competitive Priorities and except where noted the action steps below apply to all partner districts.

- **Action1.a: Establish project infrastructure.** Establish project management processes and

- **Action 1.b: Engage schools in Broward and Palm Beach Counties.** Finalize school eligibility. Involve district in activating and informing school leaders. Uncover success factors and refine plans to address each factor, such as the fit of SunBay TPD schedule to district schedules for teachers. Where needed by IRB, gain consent.

- **Action 1.c: Refine SunBay materials and district plans.** Gather all “framework” documents in the districts, such as pacing guides, additional specification of CCSSM standards, and teaching practice guidelines. Adjust documents and SunBay modules for strong alignment and to incorporate the latest information about the PARCC assessment. Test SunBay materials on variants of technology found in the districts and address issues.

- **Action 1.d: Provide teacher professional development.** Introduce all teachers to SunBay. Prepare them to teach each module. Engage teachers in reflection during and after each module. Enroll teachers in certificate and master’s TPD to build district capacity.

- **Action 1.e: Implement SunBay Mathematics in grade 6-8 classrooms.** Print paper materials. Provision user logins for digital materials to teachers and students. Confirm scheduled dates for classroom use of each module with all teachers. Track and monitor implementation by analyzing data as SunBay is used. Use analysis to identify teachers needing further support.

- **Action 1.f: Iterative improvement.** Deploy a project-wide bug and improvement/suggestion database to collect needed improvements from all partners (including evaluators). Also enable teachers and schools to contribute suggestions. Review and prioritize improvements monthly. Implement schedule of releasing improved modules to fit RCT and cost-effectiveness timetables, with at least two releases per year.

**Goal 2: Investigate models of cost-effectiveness and effective practices.** This goal, which builds on Goal 1, is central to Competitive Priority 1 and will occur only in schools in the cost-effectiveness and scaling study.

- **Action 2.a: Recruit additional CP schools.** 9 schools will be recruited in year 1 and at least another 9 in year 3, using established SRI recruitment procedures. We will prioritize the
recruitment of schools that represent the state’s diversity and that serve high-need students.

- **Action 2.b: Design variations in inputs.** Design variations in technology, curriculum, and TPD aimed at reducing cost and/or increasing productivity.

- **Action 2.c: Implement and analyze first-year comparisons.** Assemble a first set of variations in the inputs into testable SunBay modules. From Actions 1.d and 1.e, teachers will be prepared and the modules will be implemented in our cost-effectiveness and scaling districts. Data on implementation, costs, productivity, and proximal outcomes will be collected. Data will be analyzed; implications captured from Action 1.f.

- **Action 2.d: Implement and analyze second-year comparisons.** Assemble a set of improved variations into testable SunBay modules, implemented and analyzed, as in Action 2.c. Results will be used to iteratively refine the next year’s set of variations.

- **Action 2.e: Synthesize findings and implement final.** Review cumulative data and experience. Design best-case model with support for adaptation. Implement in schools.

- **Action 2.f: Collect cost data; build analytic model.** Collect data on the CP schools and the RCT districts. After RCT is complete, create a model of estimated outcomes on PARCC. Compare gains and costs across studies, as described in the section on Competitive Priority 1.

**Goal 3: Codify and refine the modular program for enhancing the effectiveness of middle school mathematics teachers.** This goal is central to Competitive Priority 2.

- **Action 3.a: Compare and synergize modular approaches.** Gather documentation on similarities and differences in how USFSP, and Lastinger Center provide TPD with technology in a modular format. Revisit and expand our literature review on effective TPD. Identify key elements of best practices. Codify them in a TPD Guiding Framework.

- **Action 3.b: Evaluate different forms of the practice.** Use the TPD Guiding Framework as the basis for TPD in the RCT (TPD by the Lastinger Center) and CP schools (TPD by USFSP). Gather implementation data on differences in TPD implementation. Analyze data to understand critical components and adaptability needs. Improve the framework.

- **Action 3.c: Develop toolkits and training.** Develop a toolkit and training to enable
additional TPD institutions and TPD leaders to deliver the framework in new settings (see Competitive Priority 2). The toolkit and training will be pilot-tested with new teacher leaders in the cost-effectiveness study as part of Action 2.d. Data gathered will be used to improve the toolkit, training, and framework.

- **Action 3.d: Assess adaptability and replicability in new districts and regions.** Within the project period, we expect demand for SunBay Mathematics beyond the Florida districts. For example, two California school districts and a district in the Shenandoah Valley are now interested in implementing SunBay. As new districts in Florida or beyond use SunBay, the adaptability and replicability of the toolkit and training will be investigated.

**Goal 4: Conduct an evaluation that meets WWC evidence standards without reservations.** The plan and actions for this goal are in Section D, Evaluation.

**Goal 5. Strengthen our capacity to expand the program upon achieving favorable outcomes, both throughout Florida and to the nation.**

- **Action 5.a: Finalize matching funds.** This proposal includes a match from the Helios Foundation (see letter of support). The Bill & Melinda Gates Foundation and Progress Energy have supported SunBay in the past and will be approached. Top SRI, UF, and USFSP leaders will be dedicated to securing the rest of the match in the first 6 months.

- **Action 5.b: Establish the St. Petersburg operations center.** The Center for Digital Learning in SRI’s Florida office will be the operations center for SunBay expansion in Florida and nationwide. Operating procedures will be codified (see also Action 5.c), and staff will be trained specifically for this purpose—to recruit schools, provide services, manage logistics, support implementations, collect feedback, and to conduct liaison with partners.

- **Action 5.c: Coordinate scale and sustainability team.** In the scaling plan, SRI will provide materials and technology, and partner universities provide TPD. Leaders from each participating partner will meet throughout the project to refine plans for further scale-up and sustainability, create joint operating procedures, and align incentives across institutions, monitor progress, and give guidance the implementation leaders at each partner institution.
• **Action 5.d: Add districts.** The partner organizations will coordinate to recruit additional Florida and national districts as the project progresses and will use the best available SunBay materials at the time they want to begin, which could be before a final version exists.

• **Action 5.e: Synthesize results of cost-effectiveness and RCT studies.** On the basis of all the feedback and data collected, including the results from Competitive Priorities 1 and 2, an ultimate design for SunBay Mathematics will be identified and readied for use.

**B.5 Competitive Priority 1: Improving Cost-Effectiveness and Productivity**

We will achieve the competitive priority of improved outcomes with reduced costs through our actions in Goal 2 [area (c) of the Competitive Priority]. The RCT will evaluate student outcomes rigorously. However, conducting an RCT adds implementation costs beyond what is normal. In year 1 we will have 9 CP schools in the exploratory cost-effectiveness study. We will investigate cost reduction and productivity by assigning teachers to variations in implementation. We will vary content delivery for one set of teachers, shifting more of the SunBay materials from print to an online format and provide web-based tools for teachers to add and adapt content as necessary. We will vary technology for another set, decreasing hardware needs from one tablet for every student to one tablet for every three students, saving money and strengthening the collaborative learning approach. We will vary TPD delivery for another set, adding additional online features that support teachers’ individual learning and collective discussion of their practices, enabling facilitators to work online more often (which saves substantial travel costs). We will use student achievement data, teacher knowledge measures, interviews and observation data to analyze our results in different implementation variations, and seek iteratively to improve each variation. In year 2 we will combine promising variations, and again analyze data and iteratively improve. In year 3 we will deploy SunBay in the 9 existing and in 9 new CP schools using the most cost-effective variation.

After making cost reduction and productivity improvements, we will combine data across the RCT and CP schools to explore whether the positive effect of SunBay on student outcomes is
maintained between the settings. We will gather more detailed information on costs in the RCT control condition (a “typical” condition), the RCT treatment condition (expensive implementation of SunBay) in collaboration with our evaluator, and in the CP schools.

We expect increased outcomes and increased costs in the RCT treatment condition relative to the control condition (Exhibit 5). In the CP schools, costs should decrease below those in the control condition while higher outcomes are maintained. In the cost comparison between the control and the basic SunBay condition, we assume that schools are acquiring some technology for students regardless of whether they are participating in SunBay (e.g., they need more technology for administering the PARCC assessments and are purchasing technology so students can practice using technology in math classes). The cost-effectiveness track, however, reduces the technology needed for math classes by aggressively implementing research-based collaborative learning techniques. These techniques will use game controllers so that three students can interact with one device, adding about  per student.

Initially, SunBay Mathematics will be a supplement to school district content costs and thus add  of materials (schools normally spend  per student in photocopying costs). As districts proceed with SunBay, we expect they will adapt materials using the SunBay platform instead of photocopying. Further, by shifting more SunBay workbooks from print to online, we will reduce supplemental cost by half. Initially, SunBay will add to TPD costs because of overlaps between district-mandated and SunBay-specific TPD sessions, requiring more TPD. In the CP schools, we will shift to more intensive but lower cost online TPD. All TPD costs assume each teacher serves 100 students and thus are reported as teacher cost per student.

The net result is an estimated per-student cost of  in the RCT condition and  for schools that implement the cost-effective version, which will be available at the end of this project. These estimates will be improved by the detailed data the project will collect both for outcomes and for costs in all three conditions. Our exploratory analytic technique will be to match schools in the control and cost effectiveness conditions on measures of prior achievement
using propensity score analysis and to compare matched schools on mean PARCC assessment scores. We will likewise compare matched schools on costs. The estimated costs are expected “ongoing costs.” The one-time cost of this validation project is [redacted]. If the program eventually reached 150,000 students per year for 5 years, 750,000 students would experience the program. The one–times costs thus add roughly [redacted] to the ongoing cost per student per year.

**Exhibit 5: Expected outcomes and costs in the RCT and cost-effectiveness study**

<table>
<thead>
<tr>
<th></th>
<th>Control RCT</th>
<th>SunBay RCT</th>
<th>SunBay cost-effectiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected effect size, per year</td>
<td>None</td>
<td>.20</td>
<td>.20 or higher</td>
</tr>
<tr>
<td>Hardware expense, $ each year (assumes [redacted] per tablet over 3 years)</td>
<td>Low collaboration</td>
<td>Medium collaboration</td>
<td>High collaboration</td>
</tr>
<tr>
<td>Game controller cost per student</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Core content costs, per student</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supplementary content costs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPD cost per student</td>
<td>Less intensive</td>
<td>More intensive</td>
<td>More intensive but more online</td>
</tr>
<tr>
<td>Estimated total cost per student</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**B.6 Competitive Priority 2: Enabling Broad Adoption of Effective Practices**

It is difficult for teachers to adopt the new practices required by the CCSSM. Through steady application of SRI’s Discipline of Innovation process over more than 5 years, we have found that modules provide teachers with a better opportunity to practice and reflect on the new pedagogical techniques than traditional PD. As the result of continuous improvement, SunBay’s modular structure now focuses teachers attention on only one area of content and one pedagogical practice a time, with ample support to try them in the classroom and reflect on their experience with the support of peers and coaches. In contrast, many programs attempt to front-load a full year’s worth of content and pedagogical techniques into an intensive summer workshop; we have shifted away from this because this approach has not worked for us.

Now we need to formalize and codify the TPD practices because as SunBay expands nationally, new university or other partners will provide TPD. This is necessary because (a) local universities or other TPD providers have credibility and relationships in their own states, and (b) TPD always involves some face-to-face contact and local providers have reduced travel costs.
Further, local providers are well positioned to design local adaptations. Over time, we have developed capacity with a cadre of teacher leaders in Texas, two universities in Florida, and a center in England. Codifying TPD practices will allow further expansion at high quality.

Our PD partners in this proposal, the Lastinger Center and USFSP, have their own successful modular site-based TPD approaches with very similar philosophies. As stated under Goal 3, under this competitive priority we will identify the common practices across different settings, and practices identified in a literature review of related work. We plan parallel implementations of TPD during this project, one in each of the settings in the logic model, and in addition to the variations we introduce in the cost effectiveness study, we expect natural variation in TPD implementation. We will use data from formative and summative evaluations to perform a cross-cutting analysis and synthesis of best practices. The results will be codified as an online toolkit for TPD leaders. The SunBay team has already begun a beta version of this toolkit that includes TPD agendas, sample videos of student work, exercises for teachers and discussion guides, walk-throughs and exercises with SunBay materials, structured opportunities to practice pedagogical moves and reflect with peers, end-of-course assessments, rubrics, and course evaluation forms. We will develop a program that trains teachers to use this toolkit and to become certified providers of SunBay TPD. Teachers who complete a master’s degree at one of our participating universities will also be eligible to be certified providers so they can spread SunBay in their schools. Per the Actions in Goal 5, we will commit to using this toolkit and training as SunBay expands, for example, to additional Florida districts and to new states. As SunBay use expands, we will collect data in these new settings about necessary and useful adaptations to the toolkit and training. Usage data from within the system and results on assessment measures will be included in our analysis of replicability, included in our improvement database, and used to guide further improvement.

C. Management Plan

SRI’s well-established project management practices clarify responsibilities, foster frequent
communication to ensure that all team members are working toward project objectives, and promote careful monitoring and feedback. Described here are the leadership roles and responsibilities, a schedule for the Goals and Actions, and performance metrics. For clarity we have included a graphic organization chart in Appendix J, section J.1.

C.1 Roles and Responsibilities

The **project director** will lead the overall project design and execution, including managing all staff and partners to meet goals within budget and on time. He will work with SRI’s internal operations team who will provide financial and administrative services for successful execution. Working closely with the project director, the **project supervisor** will see that the project goals are being addressed consistently and coherently and that that the SunBay theory and vision are maintained as iterative improvements are made. If any barriers to meeting project goals arise, the project supervisor has the authority to reallocate SRI resources as necessary. The project supervisor will also lead activities in Goal 5, related to further scale and sustainability. An **operations manager** will lead all of SRI’s implementation activities in Florida. An SRI **data manager** will oversee all data collection.

UF will appoint the **RCT implementation leader** and USFSP will appoint the **CP schools implementation leader** to work with the operations manager and district coordinators. Each partner district will appoint (and where necessary hire) a **district coordinator** who will align all factors in the district necessary for smooth implementation, most likely with the support of a **technology support coordinator** (dedicated to the project to address any technology issues in a timely manner) and a **mathematics coordinator** (to mutually align SunBay and the pacing guide, align TPD plans, etc.). The **TPD facilitators** already at UF and USFSP will lead the intensive TPD each summer and the ongoing sessions during the year. In our cost-effectiveness study, we will vary the types and numbers of PD facilitators to investigate cost reduction. The **evaluation director** will lead the RCT (Section E).

An **improvement manager** will be responsible for gathering necessary and suggested
improvements to all three SunBay components (technology, curriculum, and TPD) and will manage the schedule, resources, and prioritization of improvements. The design lead will lead the design and development of all improvements to materials and technology.

Six cross-cutting teams will coordinate the aspects of work that require input and performance across institutions and areas of expertise. These teams include Leadership (LeadT), Evaluation (EvalT), RCT Implementation (RCT-T), Competitive Priorities (CPS-T), SunBay Improvement (Sun-T), and Scale and Sustainability (Scale-T) (a table with more detail is in Appendix J, Section J.2). In-person meetings, teleconferences, and collaborative tools will be used to keep work moving forward efficiently.

**C.2 Timelines and Milestones**

Exhibit 6 lists each action or milestone in the project by year. The responsible leader and team are named. To fit the school year, actions are in spring (Sp), summer (Sm), or fall (F).

**Exhibit 6: Timeline of activities and milestones**

<table>
<thead>
<tr>
<th>Action</th>
<th>Activity or (*) Milestone</th>
<th>When</th>
<th>Leader and Team</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Year 1</td>
<td>Project Year 1 Instructions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.a</td>
<td>Kick off Project Mgt, Set Up Teams</td>
<td>Sp</td>
<td>Project director, Lead-T</td>
</tr>
<tr>
<td>1.a</td>
<td>Regular leadership meetings</td>
<td>All</td>
<td>Project director, Lead-T</td>
</tr>
<tr>
<td>1.a</td>
<td>* IRB complete</td>
<td>Sp</td>
<td>Evaluation director, Lead-T</td>
</tr>
<tr>
<td>1.b</td>
<td>Engage schools RCT</td>
<td>Sp</td>
<td>RCT implementation dir, RCT-T</td>
</tr>
<tr>
<td>1.b</td>
<td>* School data sharing agreement signed</td>
<td>Sp</td>
<td>Project director, Lead-T</td>
</tr>
<tr>
<td>3.a</td>
<td>Develop common TPD framework</td>
<td>Sp</td>
<td>Project director, Lead –T</td>
</tr>
<tr>
<td>2.a</td>
<td>Recruit schools for Competitive Priorities</td>
<td>Sp</td>
<td>Project director, Lead-T</td>
</tr>
<tr>
<td>2.b</td>
<td>Design first variations for CPS</td>
<td>Sp</td>
<td>Design lead, Sun-T</td>
</tr>
<tr>
<td>1.c</td>
<td>Refine materials, district plans for RCT</td>
<td>Sp, Sm</td>
<td>Design lead, Sun-T</td>
</tr>
<tr>
<td>4</td>
<td>Assign schools to conditions</td>
<td>Sp</td>
<td>Evaluation director, Eval-T</td>
</tr>
<tr>
<td>5.a</td>
<td>* Matching funds finalized</td>
<td>Sm</td>
<td>Project supervisor, Scale-T</td>
</tr>
<tr>
<td>1.a</td>
<td>* Complete staffing</td>
<td>Sm</td>
<td>Project director, Lead-T</td>
</tr>
<tr>
<td>5.b</td>
<td>Start St. Pete Operations Center</td>
<td>Sm</td>
<td>Project director, Operations Mgr</td>
</tr>
<tr>
<td>4</td>
<td>Provide final evaluation plan to Dept of Ed</td>
<td>Sm</td>
<td>Evaluation director, LT approves, too</td>
</tr>
<tr>
<td>1.d</td>
<td>Provide initial TPD in RCT, Cohort 1</td>
<td>Sm, F</td>
<td>Operations, RCT-T</td>
</tr>
<tr>
<td>1.d</td>
<td>Provide initial TPD in CPS, Cohort 1</td>
<td>Sm, F</td>
<td>Operations, CPS-T</td>
</tr>
<tr>
<td>2.c</td>
<td>Implement first-year variations in CPS</td>
<td>Sm, F</td>
<td>Operations, CPS-T</td>
</tr>
<tr>
<td>4</td>
<td>Baseline data collection</td>
<td>F</td>
<td>Evaluation director, Eval-T</td>
</tr>
<tr>
<td>1.e</td>
<td>Implement in classrooms in RCT</td>
<td>F</td>
<td>RCT-Implementation, RCT-T</td>
</tr>
<tr>
<td>1.f</td>
<td>Install and use database to respond to issues and prioritize improvements</td>
<td>Sm</td>
<td>Improvement mgr, design lead, RCT-T, CPS-T and SunT</td>
</tr>
<tr>
<td>1.a</td>
<td>* Year 1 metrics and performance met, * Updated management plan provided</td>
<td>F</td>
<td>Project director, Lead-T</td>
</tr>
</tbody>
</table>
### Project Year 2
2015 Calendar Year, End of 2014-15 School Year, Beginning of 2015-16 School Year

<table>
<thead>
<tr>
<th>Section</th>
<th>Activity</th>
<th>Participants</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.a</td>
<td>SunBay project annual meeting</td>
<td>Sp</td>
<td>Project director, LT</td>
</tr>
<tr>
<td>1.a</td>
<td>Regular leadership meetings</td>
<td>All</td>
<td>Project director, Lead-T</td>
</tr>
<tr>
<td>5.c</td>
<td>Scale and sustainability effort begins</td>
<td>All</td>
<td>Project supervisors, Scale-T</td>
</tr>
<tr>
<td>1.f</td>
<td>Continuous improvement of SunBay</td>
<td>All</td>
<td>Improvement mgr, SunT</td>
</tr>
<tr>
<td>1.e</td>
<td>RCT TPD and implementation continues</td>
<td>Sp</td>
<td>Operations, RCT-T</td>
</tr>
<tr>
<td>2.c</td>
<td>CPS implementation of variations continues</td>
<td>Sp</td>
<td>Operations, PCT-T</td>
</tr>
<tr>
<td>4</td>
<td>Data collection per evaluation plan, including cost data for RCT</td>
<td>Sp</td>
<td>Eval director, Eval-T</td>
</tr>
<tr>
<td>2.c</td>
<td>Analyze data, give feedback to design</td>
<td>Sp</td>
<td>Project director, Sun-T</td>
</tr>
<tr>
<td>1.f</td>
<td>Review &amp; prioritize SunBay improvements</td>
<td>Sp, Sm, F</td>
<td>Improvement mgr, Sun-T</td>
</tr>
<tr>
<td>3.b</td>
<td>Evaluate different forms of modular TPD</td>
<td>Sm</td>
<td>Project director, Sun-T</td>
</tr>
<tr>
<td>2.d</td>
<td>Design &amp; develop 2nd year variations</td>
<td>Sp, Sm</td>
<td>Design lead, Sun-T</td>
</tr>
<tr>
<td>2.d</td>
<td>Implementation TPD &amp; variations in CPS</td>
<td>F</td>
<td>Operations, CPS-T</td>
</tr>
<tr>
<td>1.d</td>
<td>TPD continues for RCT Cohort 1</td>
<td>All</td>
<td>Operations, RCT-T</td>
</tr>
<tr>
<td>1.d</td>
<td>TPD for RCT Cohort 2</td>
<td>Sm, F</td>
<td>Operations RCT-T</td>
</tr>
<tr>
<td>4</td>
<td>Data collection per evaluation plan</td>
<td>F</td>
<td>Eval director, Eval-T</td>
</tr>
<tr>
<td>1.a</td>
<td>* Year 2 metrics and performance met, * Updated management plan provided</td>
<td>F</td>
<td>Project director, Lead-T</td>
</tr>
</tbody>
</table>

### Project Year 3
2016 Calendar Year, End of 2015-16 School Year, Beginning of 2016-17 School Year

<table>
<thead>
<tr>
<th>Section</th>
<th>Activity</th>
<th>Participants</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.a</td>
<td>SunBay project annual meeting</td>
<td>Sp</td>
<td>Project director, LT</td>
</tr>
<tr>
<td>2.a</td>
<td>Recruit additional schools for CPS</td>
<td>Sp</td>
<td>Project director, Lead-T</td>
</tr>
<tr>
<td>1.a</td>
<td>Regular leadership meetings</td>
<td>All</td>
<td>Project director, Lead-T</td>
</tr>
<tr>
<td>5.c</td>
<td>Scale and Sustainability Continues</td>
<td>All</td>
<td>Project supervisor, Scale-T</td>
</tr>
<tr>
<td>5.d</td>
<td>Recruit more schools</td>
<td>All</td>
<td>Project supervisor, Scale-T</td>
</tr>
<tr>
<td>3.c</td>
<td>Develop toolkits and training, modular TPD</td>
<td>All</td>
<td>Design lead, Sun-T</td>
</tr>
<tr>
<td>1.f</td>
<td>Continuous improvement of SunBay</td>
<td>All</td>
<td>Improvement mgr, SunT</td>
</tr>
<tr>
<td>1.e</td>
<td>RCT implementation continues</td>
<td>Sp</td>
<td>Operations, RCT-T</td>
</tr>
<tr>
<td>4</td>
<td>Data collected per evaluation plan</td>
<td>Sp</td>
<td>Eval director, Eval-T</td>
</tr>
<tr>
<td>2.d</td>
<td>CPS implementation of variations continues</td>
<td>Sp</td>
<td>Operations, PCT-T</td>
</tr>
<tr>
<td>4</td>
<td>* All RCT data collected</td>
<td>Sp, Sm</td>
<td>Eval director, Eval-T</td>
</tr>
<tr>
<td>2.d</td>
<td>Analyze data, give feedback to design</td>
<td>Sp</td>
<td>Project director, Sun-T</td>
</tr>
<tr>
<td>2.e</td>
<td>Define final cost-effectiveness approach</td>
<td>Y3, Sm</td>
<td>Design lead, Sun-T</td>
</tr>
<tr>
<td>1.f</td>
<td>Develop final cost-effective approach</td>
<td>Y3, Sm</td>
<td>Improvement mgr, Sun-T</td>
</tr>
<tr>
<td>2.e</td>
<td>Implement final approach in new schools</td>
<td>Sm, F</td>
<td>Operations, PCT-T</td>
</tr>
<tr>
<td>2.f</td>
<td>Collect cost data in CPS</td>
<td>F</td>
<td>Project director, Eval-T</td>
</tr>
<tr>
<td>3.c</td>
<td>* Practice is codified and formalized</td>
<td>Sm</td>
<td>Design lead, Sun-T</td>
</tr>
<tr>
<td>3.d</td>
<td>Assess adaptability of the practice</td>
<td>F</td>
<td>Project director, Eval-T</td>
</tr>
<tr>
<td>4</td>
<td>Analysis of RCT begins</td>
<td>F</td>
<td>Eval director, Eval-T</td>
</tr>
<tr>
<td>1.a</td>
<td>* Year 3 metrics and performance met, * Updated management plan provided</td>
<td>F</td>
<td>Project director, Lead-T</td>
</tr>
</tbody>
</table>

### Project Year 4
2017 Calendar Year, End of 2016-17 School Year

<table>
<thead>
<tr>
<th>Section</th>
<th>Activity</th>
<th>Participants</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.a</td>
<td>SunBay project annual meeting</td>
<td>Sp</td>
<td>Project director, LT</td>
</tr>
<tr>
<td>1.a</td>
<td>Regular leadership meetings</td>
<td>All</td>
<td>Project director, Lead-T</td>
</tr>
<tr>
<td>5.c</td>
<td>Scale and sustainability continue</td>
<td>All</td>
<td>Project supervisor, Scale-T</td>
</tr>
<tr>
<td>1.f</td>
<td>Continuous improvement of SunBay</td>
<td>All</td>
<td>Improvement mgr, SunT</td>
</tr>
<tr>
<td>5.d</td>
<td>Recruit more districts</td>
<td>All</td>
<td>Project supervisor, Scale-T</td>
</tr>
<tr>
<td>4</td>
<td>Analysis of RCT continues</td>
<td>Sp</td>
<td>Eval director, Eval-T</td>
</tr>
</tbody>
</table>
C.3 Metrics and Performance Targets

To track finances and implementation, we will use data from financial systems, automatically collected software usage data, and submitted reports of TPD compliance and issues with the materials. The project director will assess on-time and on-budget financial performance. The operations manager will assess whether the use of the software matches expectations and will also inspect the software data for low-fidelity implementations or for bugs and use data to suggest improvements. Issues with the materials (whether in software or print materials) will be collected in a database, and the improvement manager will track resolution of each problem and report the number and kind of outstanding issues regularly to the project director. The implementations leaders will check TPD attendance against program requirements.

In year 1, the key metrics and performance targets are completion of IRB, School Data Sharing Agreements, approval of the updated evaluation plan, completed staffing, and enrolling enough teachers and schools to meet the power analysis requirements. In year 2, they are success in initial cost reductions in the cost-effectiveness study, study attrition that still meets power analysis, acceptable response rates, and low missing data rates. In year 3, they are additional cost reductions in the cost-effectiveness study, low attrition, completed data collection in the RCT with acceptable response rates and little missing data, success in adapting the practice, and success in recruiting additional schools. In year 4, the key performance targets will be completing the analyses and reports, cost-effectiveness results, success in replicating the practice, and recruiting paying districts in Florida and the nation to SunBay.

C.3 Financial and Operating Model: Plan to Operate at a Regional Level

Our operating model during this project and beyond is for SRI to distribute the SunBay
Mathematics software and materials and for university partners to provide the TPD. During this project, Lastinger Center and USFSP are the university partners (others could be added when SunBay Mathematics expands to other states) and are organized as specified in the Management Plan. Action 5.5.6 establishes an operations center at SRI to distribute the software and materials as well as a scale and sustainability team to define joint operating procedures between SRI and universities. Our actions in addressing Competitive Priority 2 and Goal 3 create the toolkit and training to replicate SunBay TPD practices to new university partners or to scale them more widely with the existing partners.

As discussed in Section B.5, we aim to develop a financial model for providing better student outcomes at lower cost to school districts. After the project ends, SunBay will be licensed on a school-wide or district-wide basis, potentially with a publishing partner. During this project costs of providing TPD are subsidized by the grant. Later, university partners will provide TPD to schools as they do today, with schools paying the cost or the cost subsidized by local philanthropies. Per Competitive Priority 1 (addressed by Goal 2), costs will be reduced by using more online and less face-to-face TPD. SunBay will also provide a reporting service so districts can track SunBay usage in their schools and monitor student outcomes.

D. Personnel

D.1. Coverage of All Key Positions

We have an established team comprising project director Phil Vahey, project supervisor Jeremy Roschelle, evaluation director Phil Sirinides, implementation leaders Phil Poekert at UF and Vivian Fueyo at USFSP, initial district coordinators for the RCT districts (Nancy Kinard at Palm Beach and Jeanine Gendron at Broward) who will be responsible for hiring final staff, improvement manager Andrew Salsbury, design lead Jennifer Knudsen, and data manager Gucci Trinidad. The district technology and math coordinators will be identified from existing staff at each district at project launch, with replacements hired by the beginning of the 2014–15 school year if districts desire to do so. SRI in St. Petersburg will hire the operations manager (Andrew
Salsbury will cover at first). SRI is able to fill most openings within 3 months.

**D.2. Qualifications and Expertise of Project Director and Key Personnel**

Dr. Philip Vahey (project director) is SRI’s director of complex mathematics initiatives, with expertise in leading large-scale initiatives directed at rigorous mathematics content, such as the SunBay initiatives in Florida and England. Dr. Vahey holds a Learning Sciences Ph.D. from UC Berkeley and has numerous publications and presentations on effective technology use in mathematics education. **Dr. Jeremy Roschelle** (project supervisor) directs SRI’s Center for Technology in Learning. His expertise is in building and sustaining large, complex programs and in developing and scaling innovative learning technologies. Dr. Roschelle has a B.S. in computer science and cognitive science and a Ph.D. in Learning Sciences; he is an internationally recognized scholar of mathematics education. **Dr. Philip Sirinides** (evaluation director), Senior Researcher at the Consortium for Policy Research in Education, leads the evaluation an i3 scale-up project *(Reading Recovery)*, and was previously the Director of Research and Evaluation for the Pennsylvania Department of Education. **Dr. Philip Poekert** (RCT implementation lead) is the Partnership Manager for the Lastinger Center, where he coordinates the center’s statewide Florida Master Teacher Initiative. **Dr. Vivian Fueyo** (CPS implementation lead), USFSP’s Interim Vice Chancellor of Academic Affairs and former Dean of the College of Education, has led large-scale PD efforts, and has strong relationships with districts throughout Florida. Both Nancy Kinard (K-12 Mathematics Education Manager for Palm Beach County Schools) and Jeanine Gendron (Director for STEM and Instructional Resources for Broward County Public Schools) have authority to serve as the initial district coordinators, and hire district staff to fill the final positions. **Andrew Salsbury** (improvement manager) oversees product development for SRI’s Education Division. **Jennifer Knudsen** (design leader) leads SRI’s math curriculum and professional development work. **Gucci Trinidad** (data manager) organizes large-scale data collection and distributed data sharing among partners on SRI’s research studies.

**E. Project Evaluation**
The Consortium for Policy Research in Education (CPRE), a leading educational research organization, will conduct a project evaluation to validate impacts and inform improvement.

**E.1 Key Questions Addressed in the Project Evaluation and Methods Used**

The primary research question is: *What is the impact of the SunBay program on student achievement, as established by CCSSM and measured by the PARCC assessments?* The secondary research question is: *What is the impact of the SunBay program on teachers’ sophistication in analyzing student understanding and instructional decision-making?* Both questions will be answered using experimental methods. To generate further hypotheses and to inform program improvement, CPRE will also investigate variation in program impacts using quasi-experimental methods. The research question for these exploratory analyses is: *What student, teacher and school factors predict variation in program effects?*

**E.2 How Evidence Produced Will Meet WWC Evidence Standards**

CPRE will conduct a 2-year cluster randomized trial. This study is designed to meet the WWC evidence standards *without reservation*. Cluster randomized control trials are internally valid in that they produce unbiased estimates of average impacts pooled across cluster units. Attention in design has been given to keeping attrition low (through strong agreements with districts), to avoidance of confounding factors (such as contamination between conditions), to validity and reliability of measures, and to avoidance of overalignment of measures to the intervention (e.g. by using PARCC). Preliminary impacts will be tested in the 2014-15 school year. The confirmatory analysis of impact will be conducted in the 2015-16 school year, which gives teachers’ sufficient time to improve their practices before student impacts are measured.

**E.3 Evaluating at the Proposed Level of Scale: Sample and Random Assignment**

Teachers in the **treatment schools** will implement the SunBay program. Teachers in the **control schools** will not have access to the SunBay program until the experiment is over and will continue with business as usual. Teachers in both conditions will attend regular district TPD, follow district pacing guides, and administer the regular district assessments. Incentives to
teachers for data collection will be the same in both conditions to assure high response rates.

**Population.** Sixty schools will be in the study (see Power Analysis, section E.4b) drawing from Broward and Palm Beach counties. These counties both serve predominantly high-need students. In both districts, most students receive free or reduced-price lunch. The Palm Beach student population comprises 28% Hispanic and 28% African American students, with 55% of students eligible for free or reduced-price lunch; the Broward student population comprises 38% African American and 29% Hispanic students, with 54% eligible for free or reduced-price lunch. Detailed demographic data is in Appendix J, Section J.3. Both districts have committed to the inclusion of at least 85% of the schools with middle grades for potential participation (avoiding schools, for example, which are being reconstituted). (Recruitment for the CPS study will also focus on schools serving predominantly minority and high-need students).

**Random sampling and assignment procedures.** Sixty schools will be randomly assigned from a subset of eligible district schools (30 in Palm Beach, 43 in Broward). Assignment at the school level rather than teacher level is to avoid within-school treatment contamination. All study schools will remain in that condition for both years of the study. All teachers and students in grades 6–8 in the participating schools will be included in the study.

Before school assignment in spring 2014, extensive school data will be gathered through administrative records and a principal survey (see below) and will be used to create blocking variables. To increase the precision of impact estimates and to explore moderating influence of school contexts, we will identify school attributes that are represented in the sample and thought to be related to SunBay implementation. We will then randomly assign schools as appropriate within groupings of these attributes. Although we do not know in advance what variables will serve as blocks, we will collect data on likely variables such as the percentage of students in the school who achieved mathematical proficiency on existing tests, school socioeconomic status, and other major programs or initiatives that might affect SunBay implementation.

**Generating evidence of differential effectiveness.** With regard to the third research question, we are able to draw on past research about SunBay Mathematics, which suggests that
effects may arise through two pathways (Roschelle et al., 2010). In many classrooms, teachers use the materials as catalysts for better teacher-student discussions about mathematical concepts. In other classrooms, the pathway to learning appears to be teacher support for student use of SunBay materials, without a particular emphasis on student discourse. Both approaches develop students’ conceptual understanding and mathematical practices and enhance student achievement. Prior SunBay research (Roschelle et al., 2010) has also found four kinds of implementation problems in classroom with weaker outcomes: Teachers (1) fail to assign students to use the technology, (2) skip sections of the written materials, (3) focus only on mathematical procedures (despite the materials), or (4) do not give students opportunities to explain, argue, or justify. All these factors are incorporated into our implementation measures and will be examined through a combination of experimental and quasi-experimental measures; by examining this implementation data thoroughly, we will generate additional hypotheses and be able to provide guidance to the design team to inform improvements to SunBay Mathematics.

**E.4 Measures, Power Analysis, and Analysis Plan**

**a. Measures.** The pre-assignment survey of schools will be administered to all principals of participating schools to collect school-level data not readily available from the districts, such as the mathematics curriculum the school uses at each grade level, the technology available at the school, and math team configurations. The data from the survey will support blocking in our random assignment process and will be useful as potential moderators in the analyses.

The 20-minute online teacher survey will be given three times to all teachers in the study over the two evaluation years. The first survey will serve as baseline, and later surveys will be retrospective for that school year. At baseline, the mathematics teacher survey will include key variables been shown to moderate the relationship between teacher learning opportunities and strong classroom implementation, including: school culture (Louis, Marks, & Kruse, 1996; Supovitz & Poglinco, 2001); academic pressure for improvement (Goldring & Pasternack, 1994); mechanisms to support teachers, including coaching and PLCs (Bryk & Driscoll, 1988);
and support for instructional approaches (Supovitz, Sirinides, & May, 2010). The baseline teacher survey will also collect participant demographics. Questions on follow-up surveys with teachers will include scales on (1) beliefs about mathematics learning, (2) mathematical practices, (3) preparation to apply the CCSSM practice standards, (4) access and comfort with technology, and (5) school context. The surveys will contain scales (and items) from existing instruments with demonstrated reliabilities. Sources for these scales will include the Teacher Education and Learning to Teach (TELT) survey (Deng, 1995), Horizon Research’s practices survey (Smith et al., 2002), and scales from other CPRE surveys. Incentives and follow-ups will be used to get high response rates.

The TASK (Teacher Analysis of Student Knowledge) will be used as the primary outcome measure of teachers’ ability to teach to high standards in their classrooms. TASK measures teachers’ ability to analyze mathematical concepts and practices within a grade-specific content area in relation to research-based learning trajectories in mathematics and to formulate effective instructional responses to student work (Daro, Mosher, & Corcoran, 2011; Szatijn et al., 2012). TASK measures several key domains including three formative assessment processes will be the focus of this evaluation, all of which are related to the CCSSM and SunBay emphasis on conceptual understanding and mathematical practices: (1) teachers’ analysis of student thinking, (2) teachers’ learning trajectory orientation, and (3) teachers’ informed instructional decision-making. CPRE has established high reliability and validity for TASK in other research projects (Ebby, Sirinides & Supovitz, 2013).

TASKs will be administered to each participating teacher (both treatment and control) in early spring 2015 and 2016. The mathematical content will be the content of the district pacing guide that was addressed by the SunBay module in the semester (for example, ratios in grade 6, proportional relationships in grade 7, and linear equations in grade 8). If teachers are teaching multiple grade levels, they will be administered the TASK relevant to the grade in which they teach the most students. A short survey will be appended to the TASK to collect information on the materials teachers used in class to teach that portion of the pacing guide. A third TASK will
be administered in late spring 2015 in a unit not associated with the SunBay curriculum to see if teachers are transferring effective mathematical practices to non-SunBay content.

Structured classroom observations will be conducted for a sample of both treatment and control teachers. The observations will focus on how teachers use the mathematical practices associated with argumentation, explanation, and justification, particularly for engaging students and eliciting and extending student thinking. The observation protocol will be a modified version of the Instructional Quality Assessment (IQA) (Junker et al., 2006; Resnick, Matsumura, & Junker, 2006). The IQA for middle school mathematics balances depth of information and feasibility of implementation and is based on research on the types of mathematics instruction that lead to improved student achievement. The IQA results in individual teacher scores on two dimensions, one related to the rigor with which mathematical content is being taught and the other to the depth of mathematical practices. The observation will contain an additional rubric used only in the treatment classrooms to generate insights into variation in implementations and impacts.

Observers will also conduct follow-up interviews with teachers in 15 treatment classrooms purposefully selected to represent varying levels of teacher outcomes on the TASK and IQA in order to understand their implementation experiences and challenges. These interviews will follow a semi-structured protocol and will be analyzed for themes as formative feedback to SR.

In addition, CPRE will use data on teacher and student technology usage from the SunBay system. This automatically collected data will provide us with a profile of classroom technology use, including topics covered, time on task, and proportion of time when many students are using technology vs. when primarily the teacher is using technology, as well as technology usage data for all students, which will contain details about completion of activities and error rates. We will use this to examine how use may be associated with outcomes.

CPRE will also assist SRI in collecting cost data in both treatment and control schools to be used in the analyses relevant to Competitive Priority 1.

Student outcome measures. The main outcome measure will be the PARCC assessment in
the 2014–15 and 2015–16 school years, administered to all grade 6, 7, and 8 students in each study school. If Palm Beach and Broward adopt additional PARCC measures such as the interim assessments, we will collect these data as repeated measures. We will also conduct exploratory analyses of student performance on district-administered interim measures that are linked to the districts’ pacing guide and mandated in all schools. These will be more proximal in time and topic to the SunBay modules but may not reflect high standards as well as PARCC. We will also collect FCAT (Florida’s existing high-stakes test) scores as baseline data.

**IRB, data collection timeline, and data management plan.** Data from schools in both conditions will be collected only after notification and consent procedures, as approved through the partners’ IRBs. The timing of the implementation of SunBay Mathematics and administration of key study instruments is shown in Exhibit 7, below. CPRE will oversee the data collection and provide access to data for secondary analysis (see Data Management Plan in Appendix J, J.4).

<table>
<thead>
<tr>
<th>Exhibit 7: Validation of Sunbay Math evaluation timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Program Year</strong></td>
</tr>
<tr>
<td><strong>Design</strong></td>
</tr>
<tr>
<td><strong>Student cohort</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Data Collection</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Work Products</strong></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

**b. Power analysis.** For research question 1, the power to detect effects is based on a conservative “intent to treat” analysis. For our calculations, the districts provided average counts
of teachers and schools. We assume that 60 schools will participate in the study, there are 12 middle school math teachers in each school, and each teacher instructs five classes of 23 students. The minimum detectable effect (MDE) sizes reported are the smallest possible standardized effect size that can be captured with 80% power for two different Type I error rates. From published literature on intraclass correlations, we expect 23% of variation in middle school math student test scores to lie between schools (Hedges & Hedberg, 2007) and 14% to lie between teachers within schools, with prior student achievement, demographics, and other school attributes able to explain 65% of the variation (Rowan, Correnti, & Miller, 2002). Power analysis for student effects using a standard three-level cluster randomized design indicates that with 60 participating schools (30 in treatment and 30 in control), the study is sufficiently powered in year 1 to detect a MDE of 0.20 standardized mean difference between groups at .05 alpha level and 0.24 at .01. For two grade levels, student outcomes will be measured in each of 2 school years separately; the correlated observations within an individual are expected to strongly increase power for treatment comparisons by the end of the study, improving the potential MDE.

We also conducted a power analysis for teacher effects (research question 2) to determine the minimum detectable program impacts on our teacher outcome measure using a standard two-level cluster randomized design. We anticipate that 10% of variation in teachers’ outcome scores is attributable to the school, of which covariates are able to explain 50% (Sirinides & Supovitz, 2012). The analysis found that with 60 schools, the study is powered for a MDE of 0.26 using the TASK. The experiment is thus sufficiently powered for both student and teacher outcomes.

c. Analysis plan.

RQ1. What is the impact of the SunBay program on student achievement? The main outcome for student achievement in mathematics will be PARCC scores (both overall scores and score for key domains). These intent to treat analyses will be performed using hierarchical linear regression models to reflect the cluster randomized design and multilevel structure of the data (Raudenbush & Bryk, 2002). To increase the precision of estimates and statistical power, vectors
of demographic characteristics for schools, teachers, and students will be included in the analysis. Additionally, multiple years of lagged student achievement scores (including FCAT) will be aggregated and centered at the school level to reduce Level 1 variance. The statistical model for estimating program impacts on student achievement is in Appendix J, Exhibit J-5.

Our analyses will also account for other teacher and student factors. First, we recognize that teacher and student mobility may cause some amount of attrition and lead to some crossover between schools. Crossing boundaries of the experimental design can be handled within a mixed-effects model by specifying cross-classified random effects. To account for the somewhat more complex data structure, we will allow for cross-classification in which lower level units may be cross-classified by two or more higher level units. A second consideration is that schools will have varying degrees of implementation and fidelity to the evaluation design, so additional analysis will assess the local effect of receiving treatment. Using the Local Average Treatment Effect (LATE) estimator (Angrist, Imbens, & Rubin 1996; Bloom, 1984) and other “treatment on the treated” methods, we will examine the dimensions of SunBay implementation that are related to teacher and student outcomes. (The LATE estimator is the ratio of the estimated impact of randomization on outcomes and the estimated impact of randomization on treatment receipt.)

**Moderator analysis.** After investigating the main impact of student outcomes, CPRE will explore for whom and under what circumstances the treatment works. We will measure potential moderators using both existing data from the district administrator and teacher and principal surveys. Because assignment of schools to treatment and control will be blocked, the analysis will potentially identify levels of these moderating variables in which it is plausible that program impacts may be larger or smaller, or where different causal mechanisms may be at work. However, because schools are randomly assigned only to the treatment and not to blocking variables (such as prior achievement or level of poverty), factors identified as associated with variation in impacts cannot be interpreted as causing the variation in impacts. Exploratory analyses of variation in impacts by demographic categories will also be conducted.

**RQ2. What is the impact of the SunBay program on teachers’ sophistication in**
analyzing student understanding and instructional decision-making? The evaluation is also designed to test whether teachers who were assigned to receive and implement the SunBay program demonstrate greater sophistication in their approach to teaching the CCSSM. The main outcome for teachers each year will be the TASK, aligned with one of the SunBay modules. We will also answer this research question using direct classroom observations to assess the quality of instruction on a sample of teachers from treatment and control, using the IQA.

**Analysis of TASK.** Overall treatment effects on teachers will be estimated using a linear model for repeated measures of the TASK. Teacher outcomes will be regressed on an indicator of treatment condition, dummy variables for each of the time points of repeated measures, an indicator of the replacement unit content, the grade level the teacher is primarily teaching, and the interactions of treatment separately with repeated measure, unit, and grade. A random school effect will be included to partition the variance in teacher outcomes between and within schools. To account for the repeated measures and resulting lack of independence among errors, we will specify an error covariance matrix to relax the independence assumption and allow for the correlation among errors between lagged repeated measures within teacher.

We will also explore the transfer of teacher skills from the SunBay program to a third non-SunBay instructional unit. This will be tested in a separate model by adding to the base model a three-way-interaction between the treatment indicator, the indicator associated with the last SunBay unit, and the indicator for transfer unit. Interpretation of this coefficient will reveal whether program effects on teachers transferred beyond the specific content in a SunBay module.

**Instructional impact.** The structured classroom observation will directly assess the quality of instruction. Multistage random sampling on all study schools will be conducted to identify 40 schools (20 treatment and 20 control) from which four middle school math teachers within each school will be selected for observation. CPRE will select one classroom session for each teacher based on two considerations: The observed class must be the same grade level as the TASK grade level that they are administered, and logistical and scheduling considerations must allow the observer to complete multiple classroom observations in 1 day. Classroom observation data
will be statistically analyzed to test for program impacts. The analytic approach will be very similar to the analysis of teacher outcomes in that the model will include a random effect for schools (to reflect cluster randomized design) and an unstructured error covariance matrix for repeated measures within classrooms. Power analysis indicates that with 160 teachers in 40 schools, we will have moderately high power ($\Delta = 0.30$) based on the wealth of teacher- and school-level information available ($\rho = 0.70$).

**Mediation analysis.** In addition to the moderator analysis, we will explicitly connect RQ1 and RQ2 to investigate the mechanisms through which the treatment operates to achieve its effect. We will measure potential mediators and implementation fidelity (below) using records of TPD, records of materials and technology use, and data from the structured classroom observations. (The technology-based measures apply only to the treatment group). Using multilevel structural equation modeling to specify the mediation of teacher practice in treatment effect on students, we will separate the total impact on student performance into direct and indirect effects. In this way, we will conduct exploratory analysis to test the theory of action supporting SunBay: that some learning occurs through student exposure to materials and collaboration among students and other learning occurs through teacher-led discussions emphasizing conceptual understanding and mathematical argumentation. The TASK, teacher survey, and classroom observations each collect relevant information. By examining the structural pathways between treatment, teacher practices, and student performance, this study will provide a better understanding of the direct and indirect statistical linkages and suggest the most effective methods of affecting instruction for students.

**RQ3: What student, teacher and school factors predict variation in program effects?**

To address this question we will conduct repeated measures analyses of outcomes with both PARCC and FCAT scores, including prior achievement (3 to 5 years of lagged achievement scores) and posttests after completing 1 and 2 years of the program. An exploratory interrupted time series analysis will use repeated measures of student outcomes to estimate individual performance trajectories for students and schools as well as produce estimates of student- and
school-level impacts (i.e., Empirical Bayes (EB) estimates). The EB methods set parameters at the highest levels of the hierarchy to the most likely values and thus shrink individual point estimates toward their population mean, reducing the mean squared error. In reporting impact variation, we will further adjust the EB estimates to address overshrinkage (Bloom, Raudenbush, & Weiss, 2012) by constraining their sample variance to equal the model-based estimate of variance of true program effects (i.e., random treatment effect for students and schools).

To permit longitudinal analysis, individual student outcomes will be benchmarked against state population score distributions to calibrate each score with a percentile ranking. If test properties and norming data for the PARCC are not available for Florida’s student population (as are available for FCAT), percentile ranking for both tests will instead be determined based on the study sample. Either approach will allow for comparison of scores across test instruments by converting student raw scores to a relative ranking, and either is congruent with the recommendations from the IES Technical Methods Report on Using State Tests in Education Experiments (May et al., 2009).

**E.5. Outcomes of Project and Measures of Implementation**

In the first 100 days of the grant, CPRE will produce a complete evaluation plan. A first report will present preliminary findings from school year 2014-15. A second report will present confirmatory analyses for RQ1 and RQ2 and analytic insights for RQ3 based on both 2014-15 and 2015-16 school years. In addition, CPRE will produce journal articles about its findings.

In addition to these summative reports, CPRE will provide SRI with data on project implementation and estimates of impacts from both the preliminary and confirmatory years. SRI will combine these with its own analyses short-term outcomes and implementation variables, with support from CPRE, seeking to refine its program logic model and identify program improvements. Improvements can be made in the scale up activities even before the RCT is complete. Finally, CPRE will review and provide constructive feedback on SRI analyses pursuant to Competitive Priorities 1 and 2.
References


