Learning Through Performance

Project-Based Learning as a Lever for Engaging the Next-Generation Science Standards

January 2021

Authors (in alphabetical order): Deutscher, Rebecca R., Holthuis, Nicole C., Maldonado, Saúl I., Pecheone, Raymond L., Schultz, Susan E. & Wei, Ruth C.*

Curriculum Design Team Leads: Susan E. Schultz and Nicole C. Holthuis

Project Lead: Rebecca R. Deutscher

Principal Investigator: Darling-Hammond, Linda

*All authors, except for S.I. Maldonado (San Diego State University), were affiliated with the Stanford Center for Assessment, Learning & Equity at Stanford University School of Education at the time the project was underway.

Contents

Acknowledgements

The Learning Through Performance science team would like to thank Lucas Education Research of the George Lucas Educational Foundation, without whose generous support this research would not have been possible. Dr. Angela DeBarger and Kristin De Vivo were important partners in conducting the research study. Dr. Jonathan Osborne and Dr. Helen Quinn were key advisors in developing our team's understanding of the intent of the Next Generation Science Standards and implications for curriculum and assessment design. Our indispensable thought partners in developing the disciplinary language supports and Complex Instruction strategies embedded within our units included the Understanding Language team at Stanford University (Dr. Kenji Hakuta, Dr. Sara Rutherford-Quach, and Dr. Annie Kuo), James Malamut (SCALE), and Dr. Rachel Lotan. A special thank you to Anne Tweedy, a major force in revising and writing much of the curriculum. We also want to recognize the research assistants involved in collecting and organizing data for the project, including Arash Jamshidi, Paolo Martin, Preetha Menon, Pai-rou Chen, Andrew Wild, and Eduardo Muñoz-Muñoz, and countless other raters and coders. The review and feedback of BSCS and Learning in Motion were enormously helpful in getting the LTP units into their final shape. Most importantly, we want to thank all of the design teachers, implementing teachers, and pilot teachers who contributed to the design of the LTP units, gave invaluable feedback on the unit designs, and opened up their classrooms to let us conduct this research.

Introduction

The Next Generation Science Standards (NGSS) (NGSS Lead States, 2013; NRC, 2012) provide a framework for shifting classroom instruction toward supporting student learning through the application of science and engineering disciplinary practices. These practices are an integral dimension of the NGSS and are typically employed by scientists and engineers as they investigate the natural world and design solutions for real-world problems.

To help students and teachers make the shift to a practice-based instructional approach, the NGSS specify a focused set of Disciplinary Core Ideas instead of the lengthy lists of skills that had comprised earlier sets of standards. By focusing on fewer ideas, students have greater opportunities to engage with the science and engineering disciplinary practices and to make connections (referred to as Crosscutting Concepts in the NGSS) across the disciplines. Students develop deeper conceptual knowledge and higher-order thinking skills versus a more traditional siloed, fact-oriented approach where students develop a superficial understanding of the content. Students learn new information and engage with evidence to answer their own "how?" and "why?" questions, leveraging a student-centered learning approach as they make sense of the phenomena under investigation. This shift equips students not only to learn new information but also to apply it.

Making this shift requires considerable curricular and instructional supports, particularly for traditionally underserved K-12 students, who may fall further behind their more advantaged peers in meeting these more rigorous standards. Research shows, however, that there is also the potential for science instruction to increase all students' academic performance in reading, writing, and science (Brown & DiRanna, 2013; Duke et al., 2020; Feldman & Malagon, 2017). Engaging *all* students in equitable and deeper exploration demands a curriculum that 1) is organized into a coherent sequence of learning, 2) includes open-ended tasks to promote interaction between students, and 3) allows students multiple opportunities to engage with meaningful phenomena and grapple with relevant questions. Moreover, effectively implementing such a curriculum also requires supports to help teachers develop their content and pedagogical knowledge.

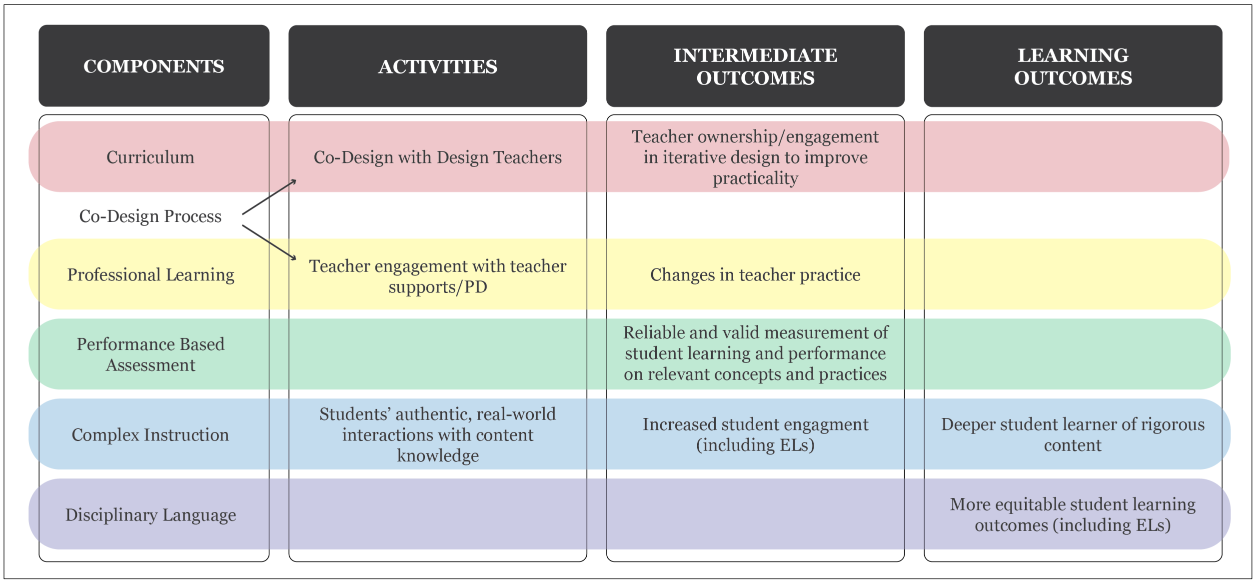
The NGSS have raised the bar for teaching, learning, and assessment in K-12 science classrooms. The goal of the Learning Through Performance (LTP) in Middle School project, funded by Lucas Education Research, was to develop, pilot, and research the efficacy of a sixth-grade science course and its related professional learning activities. The LTP science curriculum was designed by a team at the Stanford Center for Assessment, Learning and Equity (SCALE) to exemplify the use of project-based learning (PBL) as the primary instructional approach, aligned with NGSS. The LTP curriculum and professional learning were designed to support deeper learning using instructionally embedded performance-based assessments (PBA) as the primary assessment strategy, an innovative approach to measuring the deeper learning that is the goal of the shift to NGSS.[1] While PBL is an instructional approach that centers learning activity around the creation of authentic products/performances with a driving question or "need to know" that drives learning, PBA is primarily an assessment approach that allows for both group and individual demonstration of key learning outcomes. The curriculum's PBL and PBA components are detailed in the next section.

---

[1] For more on the Stanford Center for Assessment, Learning, and Equity's (SCALE's) work in science, see http://scienceeducation.stanford.edu. The website has a growing sample of free curriculum materials available.

The logic model of the LTP program (see Figure 1) proposes a set of hypotheses about how different program components will impact valued outcomes:

1. Practitioner and researcher co-development of the PBL units supports greater teacher ownership and engagement, practicality, usefulness, and developmental appropriateness of the curriculum, and university researchers and practitioners learn from each other through a reciprocal design, piloting, feedback, and revision process.
2. Designing units around PBL that incorporate Complex Instruction groupwork routines (Cohen, Lotan, Scarloss & Arellano, 1999) improves student participation; integrating performance-based assessments into PBL units supports more reliable and valid measurement of individual student learning; and embedding disciplinary language development strategies provides explicit supports for emerging biliterate students (English Learners), resulting in greater student engagement and more equitable participation and access to the curriculum. Adding these features to the design of the curriculum units are theorized to support student engagement, equitable access and participation, and supporting and measuring rigorous learning more validly.
3. Providing professional learning supports to develop teachers' pedagogical repertoire and deepen science content knowledge enables teachers to create the student-centered learning environments. These supports make good on the intent of the curriculum to elicit student engagement and deeper learning of rigorous content through authentic, real-world applications of disciplinary knowledge, concepts, and practices.



COMPONENTS	ACTIVTIES	INTERMEDIATE OUTCOMES	LEARNING OUTCOMES
Curriculum	Co-Design with Design Teachers	Teacher ownership/engagement in iterative design to improve practicality	
Co-Design Process			
Professional Learning	Teacher engagement with teacher supports/PD	Changes in teacher practice	
Performance Based Assessment		Reliable and valid measurement of student learning and performance on relevant concepts and practices	
Complex Instruction	Students' authentic, real-world interactions with content knowledge	Increased student engagment (including ELs)	Deeper student learner of rigorous content
Disciplinary Language			More equitable student learning outcomes (including ELs)

Figure 1: Logic Model Representing Theory of Action for Learning Through Performance Program

Three research questions emerged from this set of hypotheses to guide the investigation of LTP program efficacy. The research questions for the LTP efficacy study focus on how teachers-as-co-designers influenced the quality of their assignments and instruction, and how a course designed around project-based learning and performance-based assessments impacted student engagement, learning and performance. While the full LTP program and efficacy study included a math component, this report will focus only on the results for the science curriculum.

Addressing the first research question, ***How does participation as co-designers and implementers of a performance-based course influence the quality of teachers' instruction and assignments?***, study findings show that LTP science teachers who also co-designed the units improved in their ability to facilitate groupwork, to support student participation in the Learning Tasks and contribution to the Group Culminating Projects, and to manage the use of technology in their classes. Results also show that LTP science teachers' classroom activities more often involved real-world, hands-on application of science. They also increased their use of language-rich assignments and discussion strategies that require students to explain their ideas and that support their understanding of content and complete tasks; and grew in their use of performance assessment strategies (e.g., using rubrics to evaluate and provide feedback on student work).

Study findings addressing the second research question, ***How does a course that is designed around a series of performance-based assessments impact student engagement in learning?***, show that LTP students performed significantly better than a comparison group on a pre- and post-assessment designed to measure levels of engagement with the science practices. Results also indicate that when LTP science students engaged in groupwork, they were more academically engaged (e.g., manipulating materials, talking about their tasks with peers, doing projects, and making presentations) than those who were not part of these classrooms. Finally, LTP science students reported that their classroom assignments were more interesting, challenging, worthwhile, and enjoyable.

Finally, addressing the third research question, ***How does a course that is designed around a series of performance-based assessments impact student learning and performance?***, findings show that LTP science students performed significantly better than a comparison group on a pre- and post-assessment designed to measure science practices and crosscutting concepts. Additionally, student achievement in participating and nonparticipating classrooms was compared using a matched propensity score design. LTP science students outperformed matched students on the Smarter Balanced Assessment Consortium (SBAC) tests in mathematics and English language arts (ELA).

English learner performance. In addition, LTP science students classified as English learners (ELs) outperformed matched students on the California English Language Development Test (CELDT). California used the CELDT to measure students' skills in listening, speaking, reading, and writing in English during the time of this study. These results indicate that the LTP science program, which included supports for ELs, positively impacts student learning and performance in science, and positively impacts student learning and performance in mathematics, ELA, and English-language fluency and literacy. This evidence suggests that PBL programs like LTP have the potential to support more equitable outcomes for traditionally underserved K-12 students (e.g., English learner students).

LTP Science Curriculum

In collaboration with a group of middle school science teachers, the SCALE team developed a year-long, sixth-grade science curriculum aligned with NGSS standards.[2] Each unit within the curriculum embraces

---

[2] The published LTP curriculum materials can be reviewed in more detail at the following web site: https://scienceeducation.stanford.edu/curriculum/learning-through-performance-6th-grade-curriculum

project-based learning as an instructional model, incorporates rigorous curriculum-embedded performance assessments, enhances student engagement in the learning process, and provides all students access to learning using research-based groupwork strategies and language supports for ELs. The curriculum aims for students to construct a scientific explanation to answer the overarching question, "How do humans influence the world, and how does the world influence humans?" while tackling real scientific phenomena and issues.

Development Process

The SCALE content team worked iteratively with teachers to co-design and pilot the curriculum. The goal of the iterative design process was to make each unit engaging and to provide an educational experience that would be relevant and meaningful to students' lives beyond the classroom. As part of the revision process, the SCALE team also worked with faculty from Stanford University, including Dr. Rachel Lotan, an expert on language demands and heterogeneous groupwork; Dr. Helen Quinn, chair of the NGSS Framework; and Dr. Jonathan Osborne, Chair of the NGSS Science and Engineering Practices committee.

Architecture of the Curriculum

The sixth-grade science curriculum includes five units:

- **Orientation to Groupwork:** How do we work productively in a group?
- **Energy:** How do we use and control thermal energy in a system?
- **Cells & Body Systems:** How do body systems interact with each other to communicate and collaborate?
- **Variation & Heredity:** How do the environment and genetics affect who we are and how we are similar or different?
- **Climate Change:** How do we know human activity is influencing climate, and what can we do about it?

Essential Questions and Launch Activities

Each unit within the curriculum centers on a content-focused Essential Question and a Launch Activity which hooks students' interest and provides the foundation for the content and practices in that unit. The Launch Activity also introduces students to the culminating project and the criteria used to evaluate their final project. Finally, the Launch activities were also designed to help teachers gauge students' prior knowledge on science topics and phenomena and to identify possible misconceptions. (See Appendix B for an example of the components and architecture for the PBL Energy unit.)

Individual and Group Culminating Projects

Each unit includes individual and Group Culminating Projects. The Group Culminating Projects were designed to provide student choice and involve creativity and collaboration while still demanding a high level of rigor, requiring student mastery of core science content, and providing opportunities for students to demonstrate science practices. The Individual Projects provide an opportunity for students to produce evidence that they have mastered rigorous curricular standards and can demonstrate Science and Engineering Practices. These projects are open-ended and complex, and don't necessarily have one right answer. The projects not only provide an excellent opportunity for students to gain skills and content understanding but also serve as performance-based assessments, a critical component of the LTP design. The individual and group components of the projects allow students to demonstrate mastery of both rigorous content and scientific practices. Together, the projects and assessments become integrally intertwined.

Learning Tasks and Individual Organizer

To support students in preparing for the Culminating Projects, units include 3-5 key Learning Tasks and an individual project organizer. The foci of Learning Tasks are to help students identify and internalize key concepts to be applied in the Culminating Project and to allow students to engage in Science and Engineering Practices. The objectives for the Learning Tasks include not only core disciplinary content, but also science and engineering practices as well as NGSS crosscutting concepts (i.e., all three dimensions of NGSS). Within each learning task, students work in small groups and produce a group and an individual product. These products enable the teacher to provide formative feedback to students, identify and address student misconceptions, and monitor student learning throughout the process. Each learning task also contains a mini performance assessment that relates directly to the final project. At the end of each learning task, students complete a prompt in their project organizer to synthesize their knowledge for the Culminating Projects.

Rubrics

Evaluating complex, open-ended projects creates a challenge for teachers, and it is important that both students and teachers have rigorous and reliable ways of assessing the projects. Thus, the curriculum provides rubrics for evaluating the individual projects and giving students feedback. These rubrics (see pages 18-22 for a rubric example from the Energy unit) are shared with students before they begin their projects so that they understand how their work will be evaluated. The rubrics include a Science and Engineering Practice Rubric, a Science Content Rubric, and an Oral Communications Rubric.[3]

Embedded Pedagogies and Unique Hallmarks of the Curriculum

In designing the curriculum, the SCALE team maintained a focus on project-based learning and performance assessments (see also Duke, 2016; Barlowe & Cook, 2016); complex instruction through structured groupwork (see also Bennett, 2015; Cohen & Lotan, 2013); and language development opportunities. Each of these elements is discussed below.

Project-Based Learning and Performance Assessments

The projects within each unit also serve as performance assessments. When designing the curriculum, the SCALE content team embraced many of the key design principles of project-based learning from the Buck Institute for Education (now PBLWorks; see also Larmer, Mergendoller & Boss, 2015)[4] as well as

---

[3] All rubrics embedded in the LTP curriculum can be found in the teacher units at: https://scienceeducation.stanford.edu/curriculum/learning-through-performance-6th-grade-curriculum.

Please note that these rubrics do not represent SCALE's latest design thinking around integration of 3-dimensional learning and assessment that incorporate SEPs, DCIs, and CCCs.

SCALE's most recent rubric designs are reflected in their more recently developed 6-8th grade integrated NGSS curriculum at: https://scienceeducation.stanford.edu./curriculum/stanford-ngss-integrated-curriculum-exploration-multidimensional-world

[4] PBLWorks' Seven Essential Project Design Elements can be found at: https://www.pblworks.org/what-is-pbl/gold-standard-project-design

SCALE's design principles for developing rigorous curriculum-embedded performance assessments. The projects (performance assessments) enable students to:

- Access multiple forms of information needed to successfully complete the Learning Tasks, and apply and demonstrate their knowledge in different ways;
- Actively engage in their learning by making choices and decisions demonstrating self-directed learning; and
- Reflect on their learning and make revisions based on self-assessment, peer review, and/or teacher feedback.

For example, in one Culminating Project, students work in groups to use what they learn about thermal energy transfer to help solve some real-world engineering challenges. "Clients" such as Cocina del Sol, an eco-friendly Latin American food truck company, would like a device to bake their specialty cookies using the power of the sun. Another client is conducting research on Alaskan salmon and needs gloves for its researchers in Alaska who work with salmon in very cold (8°C–14°C) streams and rivers.

Students then learn relevant concepts and acquire the scientific and engineering skills to develop prototypes, test them, and revise solutions to the problems they are trying to solve. Students are provided with a variety of ways to access information and demonstrate understanding, such as designing and conducting investigations, engaging in whole-class discussions, developing explanations and models, reading text, and conducting research. As a result, these tasks provide students with multiple opportunities to read, write, listen, and talk, providing key language development opportunities. And by the end of the unit, each group has what it needs to successfully complete its project and present it to the class (Holthuis et al., 2018).

Complex Instruction and Groupwork Unit

Because the curriculum relies heavily on groupwork, the SCALE team used a framework developed by the Program for Complex Instruction at Stanford University to inform the construction of group tasks.[5] The program, which aligns with PBL key design principles, provides practical tools to support productive and equitable groupwork by incorporating three important ideas:

1. Assign student roles to promote active learning and equitable rates of participation in groups;
2. Develop activities that are open-ended and productively "uncertain," thus replacing step-by-step procedures with multiple methods for achieving multiple solutions; and
3. Shift instruction so that teachers act more as facilitators and coaches focused on group interactions, probing and challenging student thinking, and monitoring student learning.

Through implementing the curriculum, Design Teachers (those who participated in the project from the beginning of LTP curriculum development) shared a desire to know more about creating effective student groups, how to make sure that all students contribute to completing the tasks, what to say or do when a group is not functioning, and how to deal with status issues within the group. As a result, the SCALE team designed a Groupwork Unit to introduce the curriculum. The Groupwork Unit includes five

---

[5] More information about the Program for Complex Instruction can be found at: https://complexinstruction.stanford.edu/

skill-building tasks designed to help students experience what it means to collaborate and interact effectively in a group. The five skill-building tasks are as follows:

- **Task 1: Circle Challenge** requires students to construct circles using paper pieces without talking.
- **Task 2: Role Playing** helps students learn about each role and what it means to play that role.
- **Task 3: Design Master** requires one student to create a design that is hidden from the rest of the group. The master designer then gives the other students verbal instructions so they can replicate the design without being able to see it.
- **Task 4: Off Like a Rocket** is a process to help students learn how to have a productive discussion.
- **Task 5: Practicing Groupwork Skills to Construct Scientific Explanations** asks students to apply what they have learned in the previous activities to construct a group explanation using claim, evidence, and reasoning.

Focusing on Language Development

When designing and revising the curriculum, the SCALE team focused on improving language clarity to make the content accessible for all students, including English learners. Collaboration and concurrent work with the Understanding Language Center at Stanford University (led by faculty sponsor Professor Kenji Hakuta) led the LTP science team to integrate research-based practices that support students' disciplinary language development and "language in action" in authentic classroom contexts.[6] For example, throughout each unit, pedagogical supports such as visuals, sentence stems, and modeled responses support language development for all students, but especially for emerging multilingual students, and whole-group discussions are used to support both content knowledge and language development. See Appendix A for a full description of LTP language supports and examples from the Energy unit.

Leveraging Science Practices

In line with the goals of NGSS standards, the curriculum is intended to allow students to not only be science learners, but also to take on the role of scientists and engineers. As will be discussed in the findings section below, classroom observations of the piloted curriculum showed that students were engaged in problem solving, discussing, writing, reading, designing, building, and experimenting at various points throughout the curriculum. More importantly, they were engaged in productive science conversations throughout the class. In essence, students had the opportunity to practice fundamental skills that transcend science—they were developing expertise around the use and application of science and language.

LTP Professional Learning

Successfully implementing a new curriculum requires professional learning for teachers, and the SCALE team designed learning experiences that served as opportunities to both strengthen teachers' content and pedagogical knowledge, and gather feedback from teachers to inform unit revisions.

---

[6] More information on the Understanding Language Center can be found at: https://ell.stanford.edu/about

During the professional learning sessions, the SCALE team focused on the pedagogies that they had collectively identified as critical to successful implementation of the curriculum: understanding project-based learning, managing groupwork, scaffolding tasks, supporting students' language development, promoting discussions, providing students with feedback, and assessing student work.

Professional Learning Structure

Starting in the spring of 2014, the SCALE team worked with six Design Teachers to co-develop the science curriculum. Their professional learning included a five-day summer institute to prepare teachers to successfully pilot LTP materials. Throughout the rest of the school year (2013-14) and the following school year (2014-15), the team met with teachers for one full day per month to continue providing scientific content, pedagogical content, and learning experiences with the tasks. The sessions also incorporated pedagogical supports, including ways to scaffold language for all students, particularly English learners, and ways to embed Science and Engineering Practices into their instruction. In addition, teachers provided feedback and brainstormed ideas for the new units under development.

During the 2015-16 school year, three Implementation Teachers and ten new Pilot Teachers implemented the new curriculum. Implementation Teachers served as a control group during the 2014-15 school year using their own curriculum, and Pilot Teachers implemented some or all of the LTP curriculum. This round of professional learning started with a summer institute conducted across two sections: one for Design Teachers to make revisions to existing units and one for Implementation and Pilot Teachers to a) learn more about NGSS and how to unpack the performance expectations; b) understand that project-based learning and groupwork are the foundation for the development and successful implementation of the LTP curriculum; c) engage in Learning Tasks as learners; d) debrief the tasks as teachers; and e) anticipate implementation challenges with their students.

Throughout the academic year, the SCALE team also facilitated professional learning sessions for all of the project teachers during Quarterly Meetings. Each meeting was structured to ask teachers to reflect on an instructional strategy discussed at the previous meeting, focus on a pedagogical aspect of the curriculum, gather input on the unit being taught, brainstorm strategies for dealing with challenges, review key content, engage teachers in the Learning Tasks for the next unit, and/or provide teachers with a framework for the next unit. Specific objectives for each meeting were generated based on input from the teachers and/or issues observed during the classroom visits. In addition, the SCALE team visited the Design and Implementation Teachers' classrooms eight times throughout the school year and provided periodic coaching to address teacher concerns and/or questions about implementing the curriculum.

Professional Learning Topics and Strategies

Science Content Support

Some of the professional learning sessions focused on building content knowledge because teachers held varying levels of expertise in the science content. For example, as the teachers engaged in Learning Tasks from the Energy Unit, the sessions focused some of the more challenging concepts related to thermal energy and thermal energy transfer. In addition, the science of climate change presented a challenge for many teachers. Not only had some of them not taught climate change before, but the scientific body of evidence for the causes and implications of climate change is constantly expanding. The teachers in the group who had considerable experience with science knowledge helped others grapple with the science concepts, and the SCALE team also brought in guest speakers with expertise in different scientific areas.

Experiencing Learning Tasks as Learners

When learning about the curriculum, teachers had the opportunity to engage in some of the learning experiences as students, which helped to prepare teachers to facilitate the curriculum in their classrooms and to generate feedback for curriculum development. After participating in the Learning Tasks, teachers were provided time to identify potential challenges, brainstorm strategies, plan their instruction, and discuss ideas for scaffolding the students' learning between the tasks. Many of the teachers' comments, suggestions, and feedback led to specific edits while others led to significant shifts in the curriculum design.

Modeling and Peer Facilitation

While teachers were engaged as learners, they also had the opportunity to see another teacher or SCALE team member model the curriculum. The modeling process enabled teachers to see not only the content of the curriculum being taught but also the strategies being implemented. For example, the SCALE team modeled how to lead skill builders in helping students learn how to interact within a group, and provided teachers with strategies to hold students accountable for playing their roles and following the norms. Once the Design Teachers had used the curriculum in their classrooms, they played a valuable role in modeling lessons for the Implementation and Pilot Teachers. The new teachers appreciated being able to hear the lessons learned from the Design Teachers and to ask specific implementation questions.

Analysis of Student Work

Many teachers requested support in implementing assessments and using rubrics to score student work, so some of the professional learning sessions provided teachers with opportunities to score and analyze student work collaboratively. Using the rubrics in small groups allowed teachers to discuss and clarify questions about the rubrics while also making connections between instructional moves that led to stronger student work products. For example, during one particularly valuable session, teachers brought in student Culminating Projects from the Body Systems Unit. Using the LTP content and practices rubrics, they evaluated some of the students' work. Students in one of the schools had particularly strong products. The teacher of these students described scaffolding the tasks for the students so that 1) the connection between each learning task and the Culminating Project was clearer, and 2) the teacher could provide continual feedback to help students make necessary revisions over the course of the unit. This collaboration between teachers resulted in the creation of the Project Organizer, which became a key feature within each unit.

Lessons Learned about Key Conditions for Teachers' Professional Learning

Because the design and development of the professional learning components were not addressed as formal research questions in the study, insights about supporting teachers' professional learning are presented here.

Frequency of Sessions

The amount of professional learning provided to teachers is important, and almost all of the teachers commented about the need for more meetings during the academic year. In the first year of piloting the LTP curriculum (Sept 2014-June 2015), the SCALE team involved teachers in about 13 full days of professional learning, including one Summer Institute (five days) and eight full-day professional learning days across the year. After that first pilot, the number of professional learning days during the academic year was reduced from eight to four (e.g., from monthly during the school year to quarterly) in order to test the viability of the curriculum without intensive professional learning. It became immediately apparent that four days of PD across the year was insufficient, especially as it related to preparing teachers to teach each unit, learn the (new) content of the units, understand how to administer and score performance tasks, manage groupwork effectively, support student agency, teach content effectively through PBL, and support emerging biliterate students (English learners) or students with learning disabilities. Consequently, in the 2015-16 school year, some teachers were not as equipped to implement the LTP curriculum effectively as the prior cohort who had more professional learning opportunities.

Given these experiences, the SCALE team recommends at least eight full days of PD across an academic year (on top of a four- or five-day Summer Institute) to adequately address the content and skills that teachers need in order to teach the LTP curriculum. If meeting in person that often is not possible, then a blended approach composed of four to five in-person PD sessions and some number of virtual learning sessions around certain topics could be implemented.

Integration of Curriculum Development, Implementation, and Professional Learning

The SCALE team's experience working with the Design, Implementation, and Pilot Teachers—observing their classrooms and engaging in the professional learning sessions—was invaluable for gaining a better understanding of the specific challenges associated with project-based learning, groupwork, and performance assessment. This allowed the professional learning sessions to be tailored to teachers' needs. In providing feedback on their PD experiences, teachers reported learning about critical components of the project and instructional approach including project-based learning and task design, the importance of implementation support (e.g., the value of collaborating with others, the value of coaching, etc.), and the importance of collaboratively analyzing student work. Feedback from teachers was used to inform our recommendations about the frequency, content, and form of professional learning that should accompany the curriculum under optimal conditions.

Research Methodology

The research questions for the LTP efficacy study focused on how teachers as co-designers influenced the quality of their assignments and instruction, and how a course designed around project-based learning and performance-based assessments impacted student engagement, learning, and performance:

- How does participation as co-designers and implementers of a performance-based course influence the quality of teachers' instruction and assignments?
- How does a course that is designed around a series of performance-based assessments impact student engagement in learning?
- How does a course that is designed around a series of performance-based assessments impact student learning and performance?

The curriculum and professional learning were developed in Year 1 of the study (Jan 2014-May 2014). These were studied in both Year 2 (Sept 2014-June 2015) and Year 3 (July 2015-June 2016) of the project.

Study Samples

Study Teachers

Study teachers were delineated by three main groups:

1. **Design Teachers**: These teachers had participated in the project from the beginning of LTP curriculum development (Year 1, spring 2014), piloted the LTP curriculum in Year 2 (2014-15), supported curriculum revision in summer 2015, and piloted the revised LTP curriculum in Year 3 (2015-16). During curriculum development, they also participated in baseline data collection about their instructional practices.

2. **Implementation Teachers**: These teachers participated in the project in Year 2 (2014-15) as “non-treatment” teachers and also in baseline data collection about their instructional practices and student engagement, learning, and achievement. In Year 3, they participated in piloting the revised curriculum (2015-16) and research activities to examine their instructional practices, and student engagement, learning, and achievement.

3. **Pilot Teachers**: These teachers piloted some or all the revised LTP curriculum during Year 3 (2015-16) mainly for the purpose of providing additional feedback on the units. Teachers were requested to teach at least two of the units. They did not participate in research activities, other than the instructional logs.

Table 1: Number of Teacher Participants Across Years 2 and 3 of the Study

Group	**Year 2** (Sept. 2014-June 2015)	**Year 3** (July 2015-June 2016)
Design Teachers	6	3
Implementation Teachers	4	3
Pilot Teachers	NA	10

As shown in Table 1 above, 10 teachers participated in the Design and Implementation groups in Year 1 and Year 2 of the study (six Design, four Implementation). In Year 3 of the study this number was reduced to six teachers (three in each group), while an additional 10 teachers participated in piloting the LTP science curriculum.

Teacher attrition was low in Year 1, with two teachers leaving who were then replaced for Year 2. Teacher attrition was moderate in Year 2. Two design teachers and two implementation teachers left the study (four total). Teachers left the study for a variety of reasons (some personal, such as pregnancies; and some professional, such as job changes or changes in district policies). None of the teacher who left declined to participate because they did not want to implement the LTP curriculum.

Study Students

Characteristics of students in the design teacher and implementation classes are shown in Table 2 below. All data were reported by students via survey. When reporting their race/ethnicity, students were instructed to select all options they felt were applicable.

School data are reported in Table 3, showing the number of English learners as well as the number of students receiving free or reduced-price lunch in each study year. Smarter Balanced Assessment performance data for each school are presented in Appendix D.

Table 2: Students' Self-Reported Gender and Race/Ethnicity Across Years 2 and 3 of the Study

	Design Teachers Year 2 Spring		**Design Teachers Year 3 Spring**		**Implementation Teachers Year 2 Spring**		**Implementation Teachers Year 3 Spring**	
Gender								
Male	185	51.50%	94	57.30%	71	40.30%	61	44.20%
Female	169	47.10%	70	42.70%	104	59.10%	77	55.80%
Total	354	98.60%	164	100.00%	175	99.40%	138	100.00%
Race/Ethnicity (Select All That Apply)								
Asian	127	35.40%	103	60.20%	38	21.60%	18	13.00%
Hispanic/Latino	127	35.40%	48	28.10%	75	42.60%	74	53.60%
Black / African American	29	8.10%	4	2.30%	18	10.20%	21	15.20%
Pacific Islander	19	5.30%	7	4.10%	17	9.70%	5	3.60%
White	119	33.10%	26	15.20%	56	31.80%	45	32.60%
Arabic / Middle Eastern	4	1.10%	3	1.80%	3	1.70%	3	2.20%
West Indian	2	0.60%	0	0%	1	0.60%	1	0.70%
Native American (Indian)	6	1.70%	2	1.20%	9	5.10%	7	5.10%
South Asian / East Indian	6	1.70%	1	0.60%	6	3.40%	2	1.40%
Other	10	2.80%	14	8.20%	33	18.80%	25	18.10%
Total	359	100%	171	100.00%	176	100%	138	100.00%

Table 3. Percent of English Learners and Those Qualified for Free or Reduced Lunch at Study Schools*

	Percent of Students Qualified for Free or Reduced Lunch in 2014-15	Percent of Students Qualified for Free or Reduced Lunch in 2015-16	Percent of English Learners In 2014-15	Percent of English Learners In 2015-16
Design Schools				
School A	72.5	64.6	15.9	14.7
School B	56.2	53.2	22.1	22.4
School C	27.5	34.3	11.1	12.5
School D	71.6	62.5	22.4	27.9
School E	13.9	15.4	3.2	4.9
Implementation Schools				
School F	74.0	73.2	28.5	26.0
School G	85.0	94.1	31.8	23.3
School H	48.4	45.9	18.9	16.6

*All schools are public schools.

Data Sources

The data collected and analyzed to address the study research questions included classroom observations, instructional logs, teacher and student surveys, teacher interviews, LTP science unit performance tasks, LTP science pre- and post-course assessments, Smarter Balanced Assessment Consortium Math and English Language Arts Assessments, and California's English Language Development Test.

Videotaped classroom observations were coded for Year 1-3 classrooms (see Appendix C for a summary of the video coding process and findings). Responses to instructional logs and surveys were compiled and aggregated. Teacher interviews were transcribed and analyzed for themes related to research questions. Two LTP science unit performance tasks were scored by an external panel including science teachers not part of the study[7]. The panel used the provided rubrics (see footnote 3 on p.9 of this report for information about performance task rubrics). Tasks were analyzed to produce findings about student outcomes, but also the reliability and dimensionality of the rubric scores. LTP science pre- and post-course assessments were scored by an external panel (again, including science teachers not part of the study); these scores were also monitored for reliability. Large-scale assessment data (state test scores) were used to examine the effects of students' participation in LTP on student achievement for Years 2 and 3 of the study.

---

[7] In some units, tasks were not implemented as teachers ran out of time. In these cases, the tasks were not scored.

Findings

The table below provides a summary of the findings followed by a more detailed description of findings organized by research question.

Table 4: Summary of Findings

Research Questions	Summary of Findings
How does participation as co-designers and implementers of a performance-based course influence the quality of teachers' instruction and assignments?	Teachers . . . • improved in their ability to facilitate groupwork to support student participation. • used more learning activities to involve real-world, hands-on applications. • improved in their ability to facilitate authentic, active student engagement and agency. • grew in their ability to manage the use of technology or increased the use of technology. • increased their use of language-rich assignments and providing students with equitable opportunities to learn challenging academic content. • grew in their use of performance assessment strategies to evaluate and provide feedback on student work and monitor student learning. • improved in setting a clear, connected, and coherent context for learning.
How does a course that is designed around a series of performance-based assessments impact student engagement in learning?	Students' . . . • level of engagement increased throughout the project. • passive level of engagement increased. • reported level of interest, engagement, and work habits improved over time.
How does a course that is designed around a series of performance-based assessments impact student learning and performance?	Students . . . • performed significantly higher on their science post-assessment than their pre-assessment. • performed significantly better on most of the science practices on their post-assessment compared to their pre-assessment. • using the LTP curriculum performed significantly better on both the English Language Arts and Mathematics Smarter Balanced Assessments than students who did not use the LTP curriculum. • that were ELs using the LTP curriculum performed significantly better on the California English Language Development Test (CELDT) than EL students who did use the LTP curriculum.

Research Question 1: How does participation as co-designers and implementers of a performance-based course influence the quality of teachers' instruction and assignments?

Findings from classroom observations, coding of baseline to pilot year instructional videos, and teacher and pupil surveys of instructional practices show that LTP science teachers improved in their ability to facilitate groupwork to support student participation in the Learning Tasks and contribution to the Group Culminating Projects, and in their ability to manage the use of technology in their classes. Results also show that LTP science teachers increased their use of learning activities that involved real-world, hands-on application of science; increased their use of language-rich assignments and discussion strategies that require students to explain their ideas and that support their understanding of content and complete tasks; and grew in their use of performance assessment strategies (e.g., using rubrics to evaluate and provide feedback on student work).[8]

Research Question 2: How does a course that is designed around a series of performance-based assessments impact student engagement in learning?

Findings from classroom observations (with a focus on "collegial interaction" sampling) and teacher/pupil surveys indicate that levels of student engagement significantly improved over the course of the study, particularly from Year 2 to Year 3. These shifts may be a result of the concerted efforts by the science team to work closely with teachers to generate ideas about how they may better manage groupwork, and to create a new unit, implemented at the beginning of Year 3, that provided students practice and language around the behaviors and roles they should be exhibiting during groupwork. In the professional learning sessions held during summer 2016, the SCALE team was much more explicit about presenting and modeling ways to address the challenges of groupwork.[9]

While disengagement decreased, the level of students' passive engagement (looking, listening, reading, writing, etc.) went up. Our classroom observations led the research team to hypothesize that when teachers became more skilled at managing groupwork, one unintended result was a reduction in the number of students' informal, sometimes off-task, active interactions.

Teachers also indicated in interviews that Learning Tasks and Group Culminating Projects were engaging and motivating for all students, while the Individual Culminating Projects were less engaging and motivating. The units were sometimes less successful in engaging and motivating emerging bi-literate students, students with IEPs, or those with other learning challenges.

Many teachers indicated that they saw significantly positive shifts in the way students engage in the tasks and manage their work, but were also concerned that struggling students were more likely to feel overwhelmed or frustrated. The projects were complex, and students were expected to make sense of the content instead of relying on their teachers, so it was surprising that more students were not overwhelmed.

---

[8] See also Table 6 in Appendix D.

[9] See also Table 6 in Appendix D.

Science students of Design Teachers reported greater interest, engagement, and improved work habits over time, while students of Implementation Teachers reported *decreases* on these measures. This may have to do with the need for more intensive professional learning expressed by the Implementation Teachers, who did not benefit from the full 13 days of PD that Design Teachers experienced.

Research Question 3: How does a course that is designed around a series of performance-based assessments impact student learning and performance?

Study findings show that in Year 2 of the study, 139 LTP science students performed significantly better than 145 students in a comparison group on a pre- and post-assessment designed to measure science practices (asking questions and defining problems, developing and using models, planning and carrying out investigations, analyzing and interpreting data, constructing explanations and designing solutions, and engaging in argument from evidence) and crosscutting concepts (systems and system models, cause and effect, energy and matter).[10] LTP students significantly outperformed students in the comparison group by 11 percentage points ($p<.05$). In Year 3, LTP students also made significant learning gains on the same science assessment administered before and after the course (ES=0.467, $p<.001$ for items on Part A of the assessment, and ES = 0.597, $p<.001$ for items on Part B of the assessment).

LTP science students also outperformed matched students on standardized measures of student performance, including summative assessments of mathematics, ELA, and English-language fluency and literacy. These analyses of student achievement outcomes were conducted using a propensity score matching methodology, and findings are detailed below.

Propensity Score Matching

A matched linear regression analysis was conducted to examine the impacts of LTP science participation on student performance in Years 2 and 3 of the study on standardized measures, including the SBAC Mathematics and ELA assessments, and the CELDT assessment. A propensity score matching methodology was used to create the matched samples (see Pan & Bai, 2015; Powell, Hull & Beaujean, 2020). Details of Year 3 analyses are presented here, which replicate those used in Year 2.

In Year 3, only students with SBAC end-of-year summative scale scores for ***either*** mathematics or English-language arts/literacy in 2015-16 were included in the sample of 8,445 students. The LTP Science group consisted of 347 participants (curriculum across three districts and one charter school) and 8,098 non-Science LTP participants. Of these 347 students, 57 were designated as English learners, and 87 also participated in the LTP Mathematics curriculum. The matched sample of non-LTP science students included 8,098 students across six districts and two charters. More information about the matched LTP and non-LTP samples is presented in the section detailing the propensity score matching analysis methods.

Outcomes' variables in the matched models included the following scores: a) 2015-16 SBAC Mathematics, b) 2015-16 SBAC English language arts/literacy, and c) 2015-16 CELDT assessments.

---

[10] See also Tables 8-12 in Appendix E.

Matched linear regression models of SBAC and CELDT outcomes in this study were created in the following sequential order for 27 distinctive analyses (see Appendix F for regression models):

1. unmatched baseline estimates;
2. unmatched estimates accounting for students' socioeconomic status, language background (not applicable on CELDT models), special education designation and gender;
3. unmatched estimates accounting for students' race/ethnicity;
4. unmatched estimates accounting for number of days students attended school in 2015-16;
5. unmatched estimates accounting for prior achievement on standardized tests; and
6. **matched** estimates accounting for students' socioeconomic status, language background (not applicable on CELDT models), special education designation, gender, race/ethnicity, number of days students attended school in 2015-16 as well as prior achievement on standardized tests in 2014-15 (grade five).

Propensity score matching is a statistical technique that transforms multiple observed covariates into a single estimate. Unlike unmatched models, results with propensity score adjustment techniques attribute differences in outcomes beyond associative correlations.

In this study, propensity scores are each student's probability statistic of participating in LTP science (i.e., treatment). Data points selected for this study's propensity score matching formula (i.e., treatment and comparison) were:

1. Participation in the free or reduced lunch program,
2. Designation as an English learner in 2015-16,
3. Classification as a Special Education student,
4. Gender, race/ethnicity, and number of days attending school in 2015-16,
5. Prior achievement on grade five (e.g., 2014-15) CELDT, Science CST, SBAC Math, and SBAC ELA,
6. School district or charter school identifier,
7. School-level percentage of students participating in the free or reduced lunch program, and
8. School-level percentage of students designated as English learners.

Due to non-randomization, treatment and comparison students may have differed in both observed and unobserved characteristics before matching, and such differences may explain disparities in observed test score outcomes. Propensity score matches are based on the statistical likelihood that a student who did not participate in LTP science *would have* participated in LTP science based on the characteristics of all students (both treatment and comparison). Thus, PSM is a quasi-experimental technique that offers evidence that participation in LTP science impacted differences on outcome measures.

Twenty-nine covariates were used in the propensity score procedure to create the nearest-neighbor logistic regression estimation algorithm. The ratio of treatment students matched to comparison students was selected using the one-to-three nearest neighbor algorithm in largest matching order, without sub-classification or replacement, discarding units outside of common support for both treatment and comparison students and .2 caliper definitions. Additional confirmatory analyses included algorithm procedures such as one-to-one nearest matching, not discarding students outside of common support, replacement techniques, and without caliper definitions.

Findings from Year 3 analyses (see Tables 13, 14 and 15 in Appendix F for detailed results) show that after accounting for students' socioeconomic status, language background, special education designation, gender, race/ethnicity, attendance, and prior achievement, *participation in the LTP science curriculum was associated with significant positive differences in both SBAC and CELDT scores in 2015-16.* As mentioned, a similar methodology was used to examine impacts in Year 2 of the study (the pilot year); the same pattern of significant benefit of LTP can be seen in those results as well.

More specifically, the results show that during Year 2 and Year 3, the LTP science students significantly outperformed matched students in both mathematics SBAC (Year 2 effect size = .13; Year 3 effect size = .19) and English language arts SBAC (Year 2 effect size = .10; Year 3 effect size = .08). In addition, the LTP science students outperformed matched students on the CELDT during both years (Year 2 effect size = .72; Year 3 effect size = .23).

Table 5: Impact of Participation in LTP on Student Performance (LTP vs. Comparison)

	Year Two (2014-15)		Year Three (2015-16)	
Assessment	Effect Size	Sample	Effect Size	Sample
SBAC English Language Arts	0.10*	328 LTP / 9,675 comparison	0.08*	347 LTP / 8,098 comparison
SBAC Math	0.13*		0.19**	
CELDT	0.72 *	33 LTP / 2,023 comparison	0.23*	57 LTP / 1,745 comparison

**p<.001, **P<.005.*

To help readers interpret the magnitude of these effects, they can be presented as percentage point differences between LTP student performance and the performance of the matched samples. On average, LTP students outperformed their peers on the mathematics test by 12 percentage points in Year 2 and by 18 percentage points in Year 3. LTP students outperformed their peers on the ELA assessment by 8 percentage points in Year 2 and by an average 10 percentage points in Year 3. LTP English learners' performance on CELDT was significantly greater than the comparison group English learners' performance by an average of 8 percentage points in Year 2 and an average of 28 percentage points in Year 3.

Discussion

Findings suggest that the LTP science program was effective at achieving its goals relative to student engagement, science achievement, and other non-science outcomes. They also provide an opportunity to learn important lessons about the curriculum co-design process, how curriculum implementation factors impacted student learning and performance, and how to design a professional learning program to best support teachers who are interested in implementing the LTP curriculum.

Curriculum design and design process. While a co-design approach has merit for supporting user ownership and usability of the curriculum and for creating mutual learning opportunities for researchers and practitioners, it can also pose some challenges. Several factors would have improved the results of the collaborative design process that was used, including more experienced design partners, more opportunities for curriculum review, and a clearer, more concise framework for guiding the design of the LTP science curriculum.

Design partners. Developing instructional materials like the LTP science curriculum requires design partners with advanced experience with project-based learning and/or performance task design, and design partners with a strong track record of high-quality science teaching. While schools were purposively selected to avoid outlier effects, each teacher's participation within selected schools was voluntary. As a result, some of the Design Teachers did not have content background or certification in science, and most did not have any PBL experience. This led to disparate roles and influence in the curriculum design process, with teachers playing more of an advisory role than a designer role. These collaboration challenges may be ones that designers and researchers frequently find in the field when experimenting with and researching pedagogical approaches that are not yet widespread. For example, practitioners doing innovative work may be working in atypical schools that are not ideal research sites. Ultimately, the impacts of a given program are dependent on its participants, and in the context of this study, the design teachers, despite their varied backgrounds, appeared to have a positive effect on those outcomes.

Curriculum review. External reviewers and practitioners need significant opportunities to review the draft curriculum at early stages and throughout a project. In this project, additional external review would have been helpful, despite the fact that the team identified this need early on. Post-pilot year review and revision by a science teacher on special assignment was an added helpful source of expertise for the revision process. In addition, external reviews by BSCS Science Learning and Learning in Motion were initiated after Year 2 (2014-15) implementation. This review was incredibly helpful for revising and polishing the units of study. These types of reviews, however, would have been more helpful earlier in the process.

Curriculum design framework. A clearer framework is needed for designing a curriculum that does not attempt to cover all grade-level standards. **In order to meet state requirements for approved curricula,** the LTP science curriculum was designed to cover all the grade six standards in California's integrated NGSS scope and sequence, and as a result, the amount of curriculum to get through ended up being too much for participants, with most pilot teachers unable to complete in all five units in science. Rather than trying to cover all of the science standards through PBL units, it may be better to pinpoint standards that are not covered in the units (particularly those that are less well-suited to PBL), making sure that teachers are prepared to teach that content in other ways, and focusing the design process on the standards that are best suited for PBL.

LTP curriculum effectiveness. The effectiveness of the LTP science curriculum for influencing student learning and performance is moderated by implementation and the capacity of educators to provide high-quality instruction in science. The LTP science curriculum was able to influence the kinds of assignments students were asked to engage with and, to some extent, teachers' practices, student engagement, and student learning outcomes. Broadly speaking, however, results from the research show that a change in curriculum alone is insufficient to support positive changes in teachers' practice, students' learning opportunities, and student achievement. Teachers who are ineffective in implementing their existing science programs will likely still be ineffective in implementing the LTP science curriculum. Teachers need a significant amount of professional learning to build their capacity to teach using PBL as the primary method of instruction. They also need time (at least two years of piloting, if not more) to fully adopt the instructional practices that are critical to PBL, thus pointing to the crucial role of coaching and ongoing PD during initial implementation. It may be unreasonable to expect effective implementation of the LTP curriculum in less than two years.

Professional development. The amount of professional learning provided to teachers is another moderating factor. In the first year of piloting the LTP curriculum, teachers participated in about 13 full days of professional learning, including one four-day Summer Institute and a monthly full-day professional learning session (eight sessions annually). After Year 2 implementation, the number of professional learning days was reduced from eight to four to explore the minimum number of days needed to effectively train teachers to use the LTP science curriculum. It became immediately apparent that four days of PD across the year was insufficient, especially as it related to preparing teachers to teach each unit, learn the (new) content of the units, understand how to administer and score performance tasks, manage groupwork effectively, support student agency, teach content effectively through PBL, teach topics not covered by the curriculum, and support emerging biliterate students or students with learning disabilities. Consequently, in some cases teachers were not fully prepared to implement the LTP curriculum effectively. It appears, then, that at least eight full days of PD across an academic year (on top of a four- or five-day Summer Institute) is needed to adequately address the content and skills teachers need in order to teach the LTP science curriculum. An alternative model to eight full days of PD (i.e., when it isn't feasible to implement the full PD model) is a hybrid approach consisting of four to five in-person PD sessions supplemented with virtual learning communities around certain topics.

Conclusion

Study findings showed that the LTP science curriculum and professional learning led to gains in student engagement, science learning outcomes, and on standardized math, ELA, and English language proficiency assessments for participating students (as measured by the SBAC Mathematics and ELA tests, the CELDT, a science pre- and post-assessment, classroom observations, teacher interviews, and pupil and teacher surveys). These outcomes were significantly better than those of nonparticipating students. The components of the LTP program, from co-design processes and comprehensive professional learning experiences, to built-in language development practices, equitable groupwork practices, and performance-based assessment, were combined to support teacher implementation of a PBL curriculum and associated instructional practices, resulting in demonstrated improvements in student engagement and learning. The statistically significant results for English learners reinforce the potential for PBL curricula with built-in language development strategies to support more equitable access for historically underserved students to rigorous learning. These results point to the promising ways that authentic, rigorous classwork and assessments, as part of project-based learning, can be harnessed to support rigorous and yet accessible teaching and learning of the Next Generation Science Standards.

References

Barlowe, A., & Cook, A. (2016, Spring). Putting the focus in student engagement: The benefits of performance-based assessment. *American Educator*. Washington, DC: American Federation of Teachers. https://www.aft.org/ae/spring2016/barlowe-and-cook

Bennett, T. (2015, Spring). Group work for good: Unpacking the research behind on popular classroom strategy. *American Educator*. Washington, DC: American Federation of Teachers. https://www.aft.org/ae/spring2015/bennett

Brown, Z.A., & DiRanna, K. (2013). *Equal access to content instruction for English learners: An example from science*. San Francisco, CA: WestEd.

Cohen, E.G., & Lotan, R.A. (2013). *Designing groupwork: Strategies for the heterogenous classroom* (3rd ed.). New York: Teachers College Press.

Cohen, E.G., Lotan, R.A., Scarloss. B. A. & Arellano, A. R. (1999) Complex instruction: Equity in cooperative learning classrooms, Theory Into Practice, 38:2, 80-86, https://doi/abs/10.1080/00405849909543836

Duke, N., Halvorsen, A-L., Strachan, S., Kim, J., & Konstantopoulos, S. (2020). Putting PjBL to the test: The impact of project-based learning on second-graders' social studies and literacy learning and motivation in low-SES school settings. *American Educational Research Journal.* https://doi.org/10.3102%2F0002831220929638

Duke, N. (2016, Fall). Project-based instruction: A great match for informational texts. *American Educator*. Washington, DC: American Federation of Teachers. https://www.aft.org/ae/fall2016/duke

Feldman, S., & Malagon, V. (2017). *Unlocking learning: Science as a lever for English learner equity*. Oakland, CA: Education Trust-West. https://west.edtrust.org/resource/unlocking-learning-science-lever-english-learner-equity/

Holthuis, N., Deutscher, R., Schultz, R., & Jamshidi, A. (2018, Summer). The new NGSS classroom: A curriculum framework for project-based science learning. *American Educator*. Washington, DC: American Federation of Teachers. https://files.eric.ed.gov/fulltext/EJ1182064.pdf

Larmer, J., Mergendoller, J., & Boss, S. (2015). *Project based learning: A proven approach to rigorous classroom instruction*. Alexandria, VA: ASCD.

Powell, M.G., Hull, D.M., & Beaujean, A.A. (2020). Propensity score matching for education data: Worked examples. *The Journal of Experimental Education, 88*(1), 145-164. https://doi.org/10.1080/00220973.2018.1541850

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

NGSS Lead States. (2013). *Next Generation Science Standards: For states, by states*. Washington, DC: National Academies Press.

Pan, W., & Bai, H. (eds.). (2015). *Propensity score analysis: Fundamentals and developments*. New York, NY: Guilford Press.

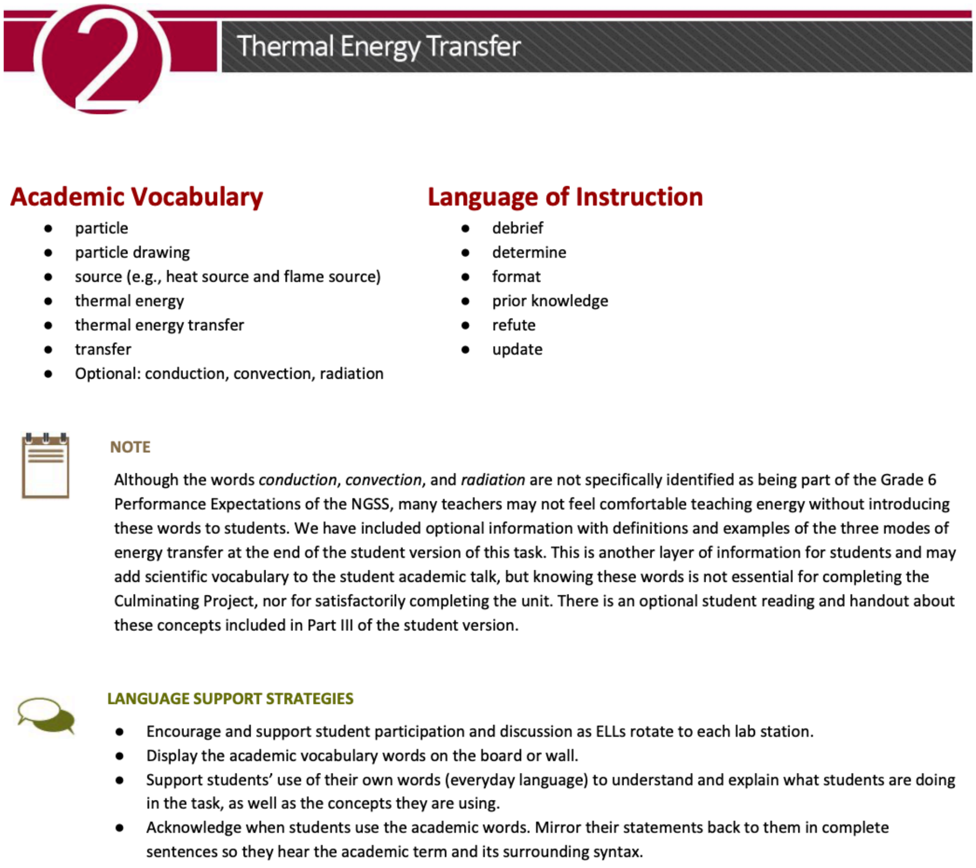
Appendices

Appendix A: Example LTP Language Development Supports – Energy Unit

Language supports that resulted from the revision process include (see examples below):

- **A one-page task card template.** The concise template provides the student with all of the learning task components at a glance, including the Essential Question; the learning objectives; the main procedures in the learning task; the materials; the rubric to be used to evaluate and provide feedback on students' performance; and the connection to the culminating task in a visually organized template. This allows students to make decisions and manage their learning process, and helps them make connections between the science practices used, the content being learned, and the Culminating Project.
- **Simplified sentence structure.** By breaking complex sentences into shorter chunks, eliminating excess nonessential words, and including vocabulary at or below grade level, the language supports all students in better understanding the purpose of the task.
- **Groupwork and language objectives**. In addition to science objectives, each Learning Task also includes groupwork/equity and language objectives.
- **Reduced number of questions**. Focused questions relate to key content, problem solving, scientific explanation, and application to relevant life situations.
- **Minimized amount of student written responses.** Learning tasks include a balance of questions for discussion and those requiring written responses.
- **Recommended science notebooks**. Due to the length of units, students need a tool to use as a reference of past concepts and practices. We added a small notebook icon next to the questions to encourage students to write in their science notebooks. This provides students opportunities to organize their thoughts; remember important concepts, skills, and practices; and record in their science notebook.
- **Aligned teacher instructions**. The teacher instructions include appropriate background material to help teachers connect practice and content. The instructions include pedagogical options, reminders of groupwork applications, the importance of class debriefing, when to provide feedback for revision, and prompts to connect with the culminating tasks.
- **Embedded language strategies.** The inclusion of specific language strategies within each unit supports emerging biliterate students' language development.
- **Language-specific objectives.** These help teachers to become explicitly aware of the language objectives to be assessed within each lesson and to emphasize the language objectives with their students.
- **Pedagogical supports.** Visuals, sentence stems, and modeled responses support language development for all students, but especially for emerging multilingual students, and whole-group discussions are used to support both content knowledge and language development.

Teacher Unit - Academic Vocabulary, Language of Instruction, and Language Support Strategies



2 Thermal Energy Transfer

Academic Vocabulary

- particle
- particle drawing
- source (e.g., heat source and flame source)
- thermal energy
- thermal energy transfer
- transfer
- Optional: conduction, convection, radiation

Language of Instruction

- debrief
- determine
- format
- prior knowledge
- refute
- update

NOTE

Although the words *conduction*, *convection*, and *radiation* are not specifically identified as being part of the Grade 6 Performance Expectations of the NGSS, many teachers may not feel comfortable teaching energy without introducing these words to students. We have included optional information with definitions and examples of the three modes of energy transfer at the end of the student version of this task. This is another layer of information for students and may add scientific vocabulary to the student academic talk, but knowing these words is not essential for completing the Culminating Project, nor for satisfactorily completing the unit. There is an optional student reading and handout about these concepts included in Part III of the student version.

LANGUAGE SUPPORT STRATEGIES

- Encourage and support student participation and discussion as ELLs rotate to each lab station.
- Display the academic vocabulary words on the board or wall.
- Support students' use of their own words (everyday language) to understand and explain what students are doing in the task, as well as the concepts they are using.
- Acknowledge when students use the academic words. Mirror their statements back to them in complete sentences so they hear the academic term and its surrounding syntax.

2

Thermal Energy Transfer

Part I • Thermal Energy Transfer Lab Stations 1–6

1. Set up lab stations with the appropriate lab equipment and lab station student direction handouts.
2. Place students in their project groups. Designate student roles and review the norms.
3. Optional: Hand out one Student Participation Observations form to each table. As you circulate among the lab stations, stamp the form as students work and discuss.
4. Ask a student volunteer in each group to read the letter from Hilton to the Science Wizard.
5. As a class or table group, take a vote to see if students think Mom or Hilton is correct about why the refrigerator warms up when the door is open. Remind students that after completing the lab stations, they may want to change their vote.
6. Ask students to write a claim in their science notebook with reasoning to answer Hilton's question. At this point, their reasoning will only be backed up by prior knowledge and may only be a guess.

LANGUAGE SUPPORT STRATEGIES

Some ELLs will need sentence frames to support the writing of their claim:

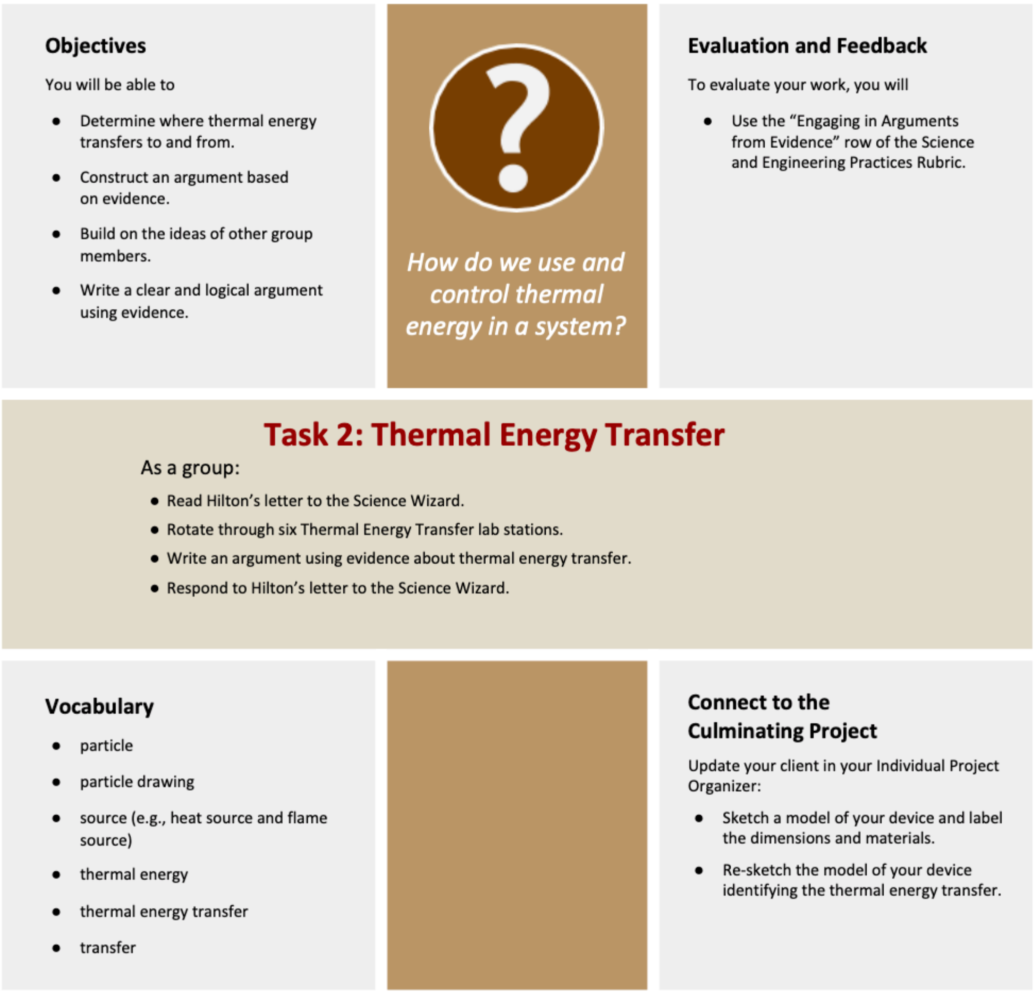
- The _____ (warm, cold) air moves into a _____ (cold, warm) space.
- I think this is true because _____. For example, _____.
- Another reason is _____, as you can see in my particle drawing. Finally, _____.

7. Review the directions for each lab station. Ask groups to rotate volunteers to read through the directions at each station.

LANGUAGE SUPPORT STRATEGIES

- Say and display (hold up when possible) the name of each of the materials for lower-level proficiency ELLs.
- When introducing Lab Station 3, stretch out the syllables in *conductometer* so they can hear the word clearly. Ask students to repeat.
- When introducing Lab Station 6, point out which side of the blanket is reflective.

Student Unit – Task Card Template



Objectives

You will be able to

- Determine where thermal energy transfers to and from.
- Construct an argument based on evidence.
- Build on the ideas of other group members.
- Write a clear and logical argument using evidence.

How do we use and control thermal energy in a system?

Evaluation and Feedback

To evaluate your work, you will

- Use the "Engaging in Arguments from Evidence" row of the Science and Engineering Practices Rubric.

Task 2: Thermal Energy Transfer

As a group:

- Read Hilton's letter to the Science Wizard.
- Rotate through six Thermal Energy Transfer lab stations.
- Write an argument using evidence about thermal energy transfer.
- Respond to Hilton's letter to the Science Wizard.

Vocabulary

- particle
- particle drawing
- source (e.g., heat source and flame source)
- thermal energy
- thermal energy transfer
- transfer

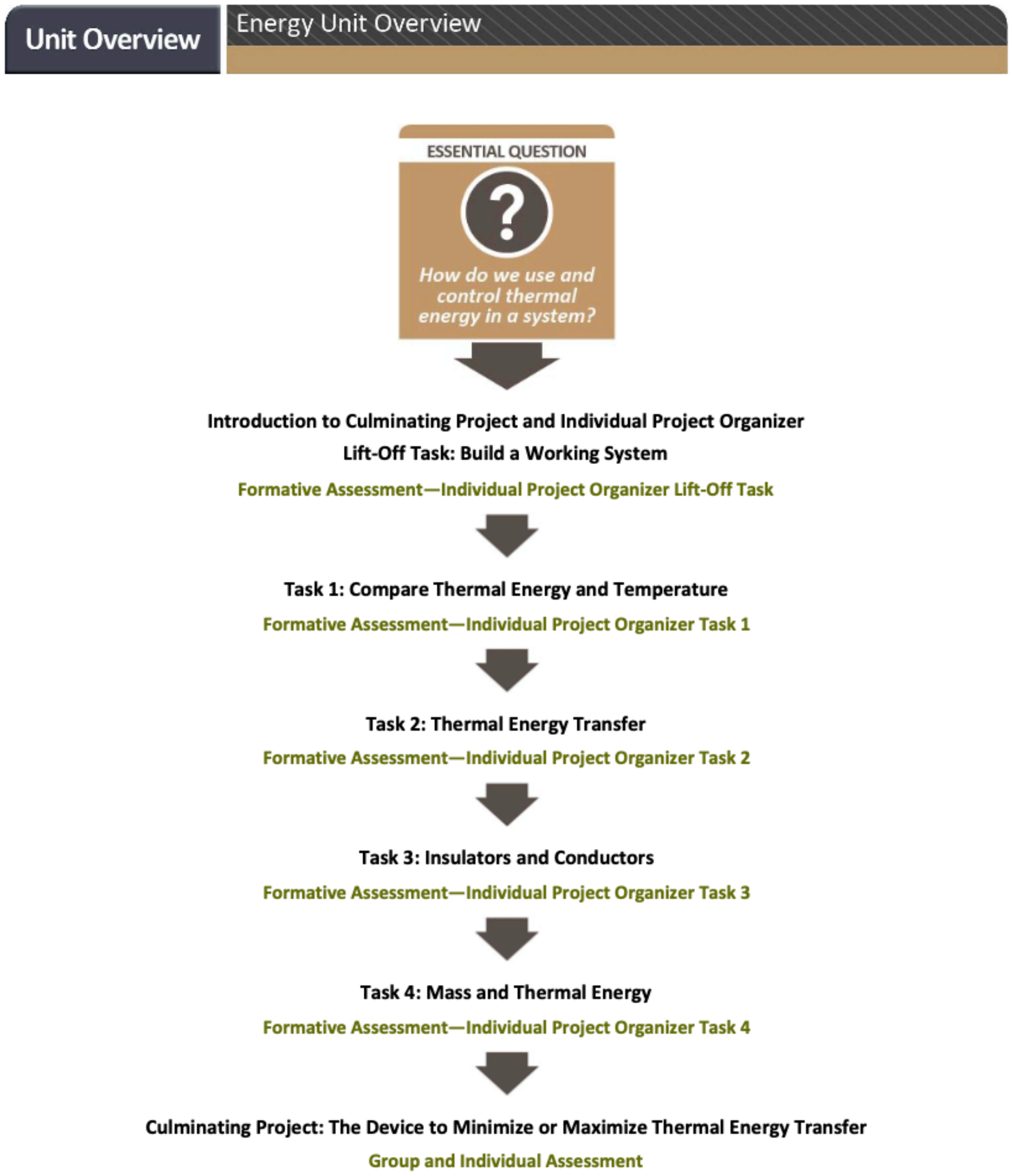
Connect to the Culminating Project

Update your client in your Individual Project Organizer:

- Sketch a model of your device and label the dimensions and materials.
- Re-sketch the model of your device identifying the thermal energy transfer.

Appendix B: LTP Unit Architecture (Energy Unit Example)

Energy Unit Overview



Unit Overview Energy Unit Overview

ESSENTIAL QUESTION

How do we use and control thermal energy in a system?

Introduction to Culminating Project and Individual Project Organizer

Lift-Off Task: Build a Working System

Formative Assessment—Individual Project Organizer Lift-Off Task

Task 1: Compare Thermal Energy and Temperature

Formative Assessment—Individual Project Organizer Task 1

Task 2: Thermal Energy Transfer

Formative Assessment—Individual Project Organizer Task 2

Task 3: Insulators and Conductors

Formative Assessment—Individual Project Organizer Task 3

Task 4: Mass and Thermal Energy

Formative Assessment—Individual Project Organizer Task 4

Culminating Project: The Device to Minimize or Maximize Thermal Energy Transfer

Group and Individual Assessment

Appendix C: Summary of Video Coding Process & Findings

To address the research question about the *influence of the LTP curriculum and PD on the quality of teachers' instruction and assignments,* lesson videos collected during classroom observations across the years were analyzed to see if changes could be observed in teachers' instructional practices over time *within* the Design Group and *within* the Implementation Group.

Video coders were trained and calibrated each year of the study, and coders deemed to be efficient and accurate continued to be leveraged year-over-year. Classroom videos were subdivided into five-minute segments to manage cognitive load and time challenges, and within each segment coders could apply each unique code only once. Twenty percent of video lessons were double coded to evaluate reliability.

Videos were coded on six primary claims about teachers' use of high-leverage PBL instructional practices:

- Claim 1: Does the LTP curriculum and professional learning help teachers get better at **facilitating authentic, active student engagement and agency**?
- Claim 2: Does the LTP curriculum and professional learning help teachers get better at **helping students actively engage in authentic Learning Tasks / activities**?
- Claim 3: Does the LTP curriculum and professional learning help teachers get better at **providing students, including those with language and academic challenges, equitable opportunities to learn rich and challenging academic content**?
- Claim 4: Does the LTP curriculum and professional learning help teachers get better at **setting a clear, connected, and coherent context for learning**?
- Claim 5: Does the LTP curriculum and professional learning help teachers get better at **formatively assessing and monitoring student learning**?
- Claim 6: Does the LTP curriculum and professional learning help teachers get better at **using technology and tools authentically**?

Overall, when comparing baseline year data (Year 1) and the year when teachers began piloting the LTP curriculum (Year 2), science Design Teachers demonstrated important changes in their use of high-leverage instructional strategies associated with PBL. In addition, when compared with other teachers (Implementation Teachers) who were not using the LTP curriculum, there were clear distinctions between the Design and Implementation Groups across most claims. These findings suggested that participating in designing and piloting the LTP curriculum supported important changes in teachers' instructional practices and assignments across the claims.

In Year 3, all teachers (Design and Implementation) who remained in the study were implementing the LTP curriculum, so rather than comparing teachers between groups, the analysis focused on *within group changes* across years (Years 1, 2, and 3 for Design Teachers, and Years 2 and 3 for Implementation Teachers), as well as changes within individual teachers from year to year.

Results also indicated that when aggregating the video coding data for teachers who were in the study from Years 1, 2, and 3, there was steady increase in teaching behaviors associated with Claims 1, 3, 4 and 5.

Among three science Design Teachers who remained in the study across all three years, there were even greater gains in Year 3 on three of the claims in particular: Claim 3, Claim 4, and Claim 5. Among three science Implementation Teachers who were in the study in Years 2 and 3, there were even more dramatic gains across all six Claims (although Claim 2 was not systematically coded in Years 2).

Appendix D: Smarter Balanced Assessment Scores by Study School

Table 6. Proportion of Students That Met or Exceeded Standard on Smarter Balanced Assessments of Math and English Language Arts in 2014-15 and 2015-16 by Study School*

	SBAC MATH		SBAC ELA	
	Percent of Students That Met or Exceeded Standard on SBAC Math 2014-15	Percent of Students That Met or Exceeded Standard on SBAC Math 2015-16	Percent of Students That Met or Exceeded Standard on SBAC ELA 2014-15	Percent of Students That Met or Exceeded Standard on SBAC ELA 2015-16
Design Schools				
School A	59	59	57	60
School B	31	40	38	45
School C	56	60	63	67
School D	17	21	24	29
School E	52	51	59	64
Implementation Schools				
School F	18	21	31	36
School G	9	14	22	38
School H	34	42	44	52

*All schools are public schools.

Appendix E: Figures and Tables of Findings

Figure 2: Number of Science Teachers That Made Adaptations (Units 1-5)

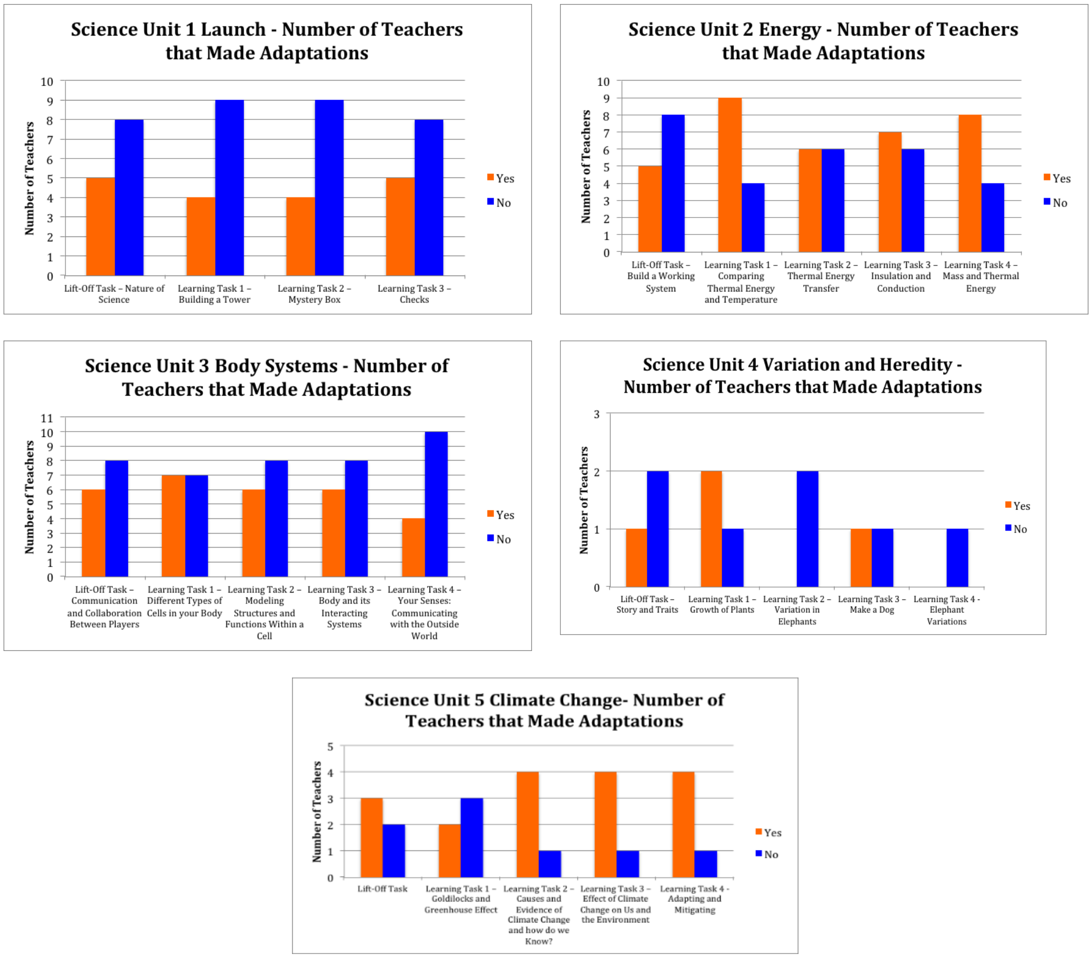


Table 7: Comparisons between Year 1 Spring and Year 3 Spring: Science Design Teachers-Pupil Survey Items on Learning Experiences With Statistically Significant Increases or Decreases in Ratings

Survey item	**Year 1, Spring 2014**	**Year 3, Spring 2016**	**Year 1, Spring 2014 vs. Year 3, Spring 2016**
	Percent "Mostly True" + "Totally True"		**Percent difference**
How true are the statements below?			
a. My teacher gives us time to explain our ideas.	61.1	74.9	**13.8**
b. In this class, we learn to correct our mistakes.	57.6	73.7	**16.1**
c. My teacher knows when the class understands and when we do not.	44.4	57.9	**13.5**
d. My teacher is nice to me when I ask questions.	56.3	69.0	**12.7**
e. In this class we have to think hard about the writing we do.	45.8	63.2	**17.4**
f. My teacher explains difficult things clearly.	44.4	54.4	**10.0**
g. My teacher wants us to share our thoughts.	51.4	64.3	**12.9**
h. My teacher in this class makes me feel that he/she really cares about me.	38.9	48.0	**9.1**
i. School work is not very enjoyable. (Do you agree?)	49.3	40.9	**-8.4**
j. My teacher pushes everybody to work hard.	50.7	69.6	**18.9**
k. My teacher wants me to explain my answers—why I think what I think.	63.2	76.0	**12.8**
How often does your science teacher do the following things? In this class, my teacher...	**Percent "Often" + "Very Often"**		**Percent difference**
a. asks me to give reasons and provide evidence for my answers.	61.8	81.3	**19.5**
b. encourages me to ask questions.	45.1	57.3	**12.2**
c. encourages me to explain concepts to other students.	27.8	53.2	**25.4**
d. encourages me to consider different scientific explanations.	39.6	64.3	**24.7**
e. provides time for me to discuss science ideas with other students.	38.9	65.5	**26.6**
How often did YOU do these things? In this class, I...	**Percent "Often" + "Very Often"**		**Percent difference**
a. debate with other students about the meaning of data.	25.0	38.0	**13.0**
b. learn from my classmates.	32.6	50.9	**18.3**
c. consider scientific explanations that are different from accepted theories.	29.2	48.5	**19.3**
d. have a say in deciding what activities I do.	21.5	35.1	**13.6**
e. write about how I solved a science task or about what I	31.9	50.9	**19.0**

am learning.			
f. learn that there are different ways to solve science tasks.	43.8	73.1	**29.3**
g. talk with my classmates about how to approach science tasks.	27.8	62.6	**34.8**
h. use science to solve real world problems.	23.6	48.0	**24.4**
i. use a computer or other technology to learn.	59.0	77.2	**18.2**
j. use the Internet for science assignments and activities.	58.3	77.8	**19.5**
l. work on science tasks in small groups.	59.0	76.6	**17.6**
m. read about science.	56.9	43.9	**-13.0**
How often did you do these things to show your teacher what you have learned? When I show my teacher what I have learned, I...	**Percent “Often” + “Very Often"**		**Percent difference**
a. discuss my ideas, approaches, and solutions with my teacher.	19.4	40.9	**21.5**
b. write sentences or paragraphs to explain my solutions.	28.5	47.4	**18.9**
c. present my ideas to an audience other than my teacher.	22.9	39.2	**16.3**
d. draw and explain the meaning of graphs or diagrams.	17.4	52.0	**34.6**
f. use a rubric to assess my own learning.	22.9	27.5	**4.6**
g. discuss my ideas, explanations, and solutions with other students.	25.7	53.2	**27.5**
h. have a chance to correct my work before I get a final grade.	41.7	62.0	**20.3**
i. use my notes.	45.8	61.4	**15.6**
j. use a computer or other technology.	49.3	69.6	**20.3**

Table 8: Science Assessment Part 1, Descriptive Statistics: Year 3 Aggregated Scores

All Design and Implementation Teachers					
	n	**Min**	**Max**	**Mean**	**SD**
Pre-Test	209	7	32	22.90	4.18
Post-Test	237	16	32	24.91	3.42

Table 9: Science Assessment Part 2, Descriptive Statistics: Year 3 Aggregated Scores

All Design and Implementation Teachers					
	n	**Min**	**Max**	**Mean**	**SD**
Pre-Test	199	0	22	13.91	3.71
Post-Test	211	8	22	15.99	3.07

Table 10: Science Assessment Part 1 and Part 2: Year 3 Aggregated Student Learning Gains

Teacher	***t* Value**	**Degrees of Freedom**	**Significance Level**	**Effect Size (Cohen's *d*)**
Part 1	-.334	177	0.000***	0.467
Part 2	-6.435	157	0.000***	0.597

***$p < 0.001$

Table 11: Year 3 Student Learning Gains on Science and Engineering Practices
(Parts 1 and 2 Are Combined)

Science and Engineering Practices	***t* Value**	**Degrees of Freedom**	**Significance Level**	**Effect Size (Cohen's *d*)**
Asking questions and defining problems	–4.226	215	0.000***	0.334
Developing and using models	–5.965	159	0.000***	0.555
Planning and carrying out investigations	–4.540	231	0.000***	0.317
Analyzing and interpreting data	–4.250	146	0.000***	0.402
Using mathematics and computational thinking	0.642	229	0.522	0.058
Constructing explanations and designing solutions	–6.828	188	0.000***	0.582
Engaging in argument from evidence	–3.665	191	0.000***	0.330

Obtaining, evaluating, and communicating information	–0.589	219	0.557	0.058

***$p < 0.001$

Table 12: Year 3 Student Learning Gains on Crosscutting Concepts (Parts 1 and 2 Are Combined)

Science and Engineering Practices	***t* Value**	**Degrees of Freedom**	**Significance Level**	**Effect Size (Cohen's *d*)**
Patterns	0.642	229	0.522	0.058
Systems and system models	–7.324	135	0.000***	0.741
Cause and effect	–5.010	230	0.000***	0.412
Energy and matter	–4.711	178	0.000***	0.424

***$p < 0.001$

Appendix F: Regression Results

Table 13: Linear Regression Results of Disaggregated Science LTP on 2015-16 SBAC Mathematics

	2015-2016 SBAC MATH Scaled Scores					
	Unmatched Results				*Matched Results*	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Fixed Effects	(Coef. *SE*)	(Coef. *SE*)	(Coef. *SE*)	(Coef. *SE*)	(Coef. *SE*)	(Coef. *SE*)
Intercept/Constant	β=2507.67*** (1.32)	β=2575.09*** (1.93)	β=2589.27*** (2.65)	β=2408.30*** (16.44)	β=507.80*** (36.83)	β=445.35*** (35.73)
Treatment						
Science LTP (Non-Science LTP omitted)	β = 31.22*** (7.03)	β = 29.88*** (6.22)	β = 22.14*** (6.05)	β = 21.73*** (5.10)	β = 19.71*** (5.39)	β = 20.50*** (2.38)
Demographic Background						
Low Socioeconomic Status (Non-low SES omitted)	-	β = -78.15*** (2.29)	β = -44.12*** (2.33)	β = -41.67*** (2.32)	β = -13.98*** (2.17)	β = -9.50*** (1.99)
ELL Designation (Non-ELL omitted)	-	β = -76.24*** (3.01)	β = -85.32*** (2.88)	β = -88.26*** (2.86)	β = -14.93*** (2.67)	β = -10.11*** (2.58)
SPED Classification (Non-LTP omitted)	-	β = -107.79*** (3.77)	β = -87.84*** (3.49)	β = -87.53*** (3.46)	β = -35.30*** (3.38)	β = -34.40*** (3.50)
Female (Male omitted)	-	β = -1.33 (2.23)	β = -1.45 (2.04)	β = -1.77 (2.02)	β = 5.55*** (1.71)	β = 4.40** (1.62)
Asian (White omitted)	-	-	β = 12.60*** (3.04)	β = 10.29*** (3.02)	β = 6.37* (2.88)	β = 1.28 (2.72)
Black (White omitted)	-	-	β = -120.02*** (4.05)	β = -117.20*** (4.03)	β = -33.71*** (3.45)	β = -37.95*** (3.28)
Hispanic (White omitted)	-	-	β = -53.54*** (3.28)	β = -52.85*** (3.25)	β = -9.14*** (2.81)	β = -11.76*** (2.55)
Native (White omitted)	-	-	β = -39.53** (14.26)	β = -38.36~ (14.13)	β = -13.12 (11.04)	β = -32.06*** (9.50)
Multi-racial (White omitted)	-	-	β = -12.45 (8.36)	β = -12.11 (8.28)	β = -12.28* (5.57)	β = -16.62** (5.51)
Race-ethnicity Missing (White omitted)	-	-	β = -27.14*** (6.99)	β = -29.10*** (6.93)	β = -22.61* (9.71)	β = -25.07* (11.20)
2015-2016 Attendance (Continuous, Mean: 172)	-	-	-	β = 1.03*** (0.09)	β = 0.63*** (0.10)	β = 0.65*** (0.10)
Grade 5 SBAC MATH (Continuous, Mean: 2494)	-	-	-	-	β = 0.73*** (0.02)	β = 0.76*** (0.02)
Grade 5 CST Science (Continuous, Mean: 355)	-	-	-	-	β = 0.27*** (0.02)	β = 0.23*** (0.02)
R^2	0.002	0.356	0.475	0.485	0.804	0.800

Key: ~ = p. < .10, * = p. < .05, ** p. < .01, *** = p. < .001

Note: These models included 347 LTP students and 8098 non-LTP students

Table 14: Linear Regression Results of Disaggregated Science LTP on 2015-16 SBAC English Language Arts

	2015-2016 SBAC ELA Scaled Scores					
	Unmatched Results					*Matched Results*
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Fixed Effects	(Coef. *SE*)	(Coef. *SE*)	(Coef. *SE*)	(Coef. *SE*)	(Coef. *SE*)	(Coef. *SE*)
Intercept/Constant	β=2513.74***	β=2563.23***	β=2580.79***	β=2448.39***	β=1006.48***	β=1053.21***
	(1.22)	(1.74)	(2.43)	(15.74)	(38.56)	(37.79)
Treatment						
Science LTP	β = 31.22***	β = 25.67***	β = 4.05*	β = 13.85*	β = 6.79	β =7.88**
(Non-Science LTP omitted)	(6.42)	(5.56)	(5.52)	(5.49)	(5.51)	(2.49)
Demographic Background						
Low Socioeconomic Status		β = -67.80***	β = -40.26***	β = -38.34***	β = -17.82***	β = -16.14***
(Non-low SES omitted)	-	(2.06)	(2.14)	(2.14)	(2.24)	(2.10)
ELL Designation		β = -82.97***	β = -91.47***	β = -91.10***	β = -24.86***	β = -28.31***
(Non-ELL omitted)	-	(2.73)	(2.66)	(2.65)	(2.82)	(2.78)
SPED Classification		β = -92.14***	β = -75.62***	β = -75.21***	β = 31.27***	β = -32.17***
(Non-LTP omitted)	-	(4.40)	(3.21)	(3.19)	(3.51)	(3.70)
Female		β = 23.23***	β = 22.91***	β = 22.70***	β = 16.02***	β = 17.47***
(Male omitted)	-	(2.01)	(1.87)	(1.86)	(1.80)	(1.75)
Asian			β = -0.82***	β = -1.02	β = 3.79	β = 6.91**
(White omitted)	-	-	(2.79)	(2.78)	(2.97)	(2.87)
Black			β = -	β = -	β = -32.43***	β = -26.02***
(White omitted)	-	-	106.70***	104.75***	(3.56)	(3.46)
Hispanic			β = -47.03***	β = -46.32***	β = -8.80**	β = -7.87**
(White omitted)	-	-	(3.00)	(2.99)	(2.91)	(2.69)
Native			β = -35.77**	β = -34.97**	β = -18.78~	β = -14.78
(White omitted)	-	-	(13.07)	(13.00)	(11.40)	(10.02)
Multi-racial			β = -15.98*	β = -15.78**	β = -16.75**	β = -18.71***
(White omitted)	-	-	(7.66)	(7.62)	(5.76)	(5.82)
Race-ethnicity Missing			β = -29.90***	β = -31.35***	β = -21.19*	β = -23.99*
(White omitted)	-	-	(6.41)	(6.38)	(10.03)	(11.82)
2015-2016 Attendance				β = 0.75***	β = 0.33***	β = 0.13
(Continuous, Mean: 172)	-	-	-	(0.09)	(0.10)	(0.11)
Grade 5 SBAC ELA					β = 0.54***	β = 0.54***
(Continuous, Mean: 2491)	-	-	-	-	(0.02)	(0.02)
Grade 5 CST Science					β = 0.33***	β = 0.34***
(Continuous, Mean: 355)	-	-	-	-	(0.02)	(0.02)
R^2	0.003	0.386	0.480	0.486	0.751	0.741

Key: ~ = p. < .10, * = p. < .05, ** p. < .01, *** = p. < .001

Note: These models included 347 LTP students and 8098 non-LTP students

Table 15: Linear Regression Results of Disaggregated Science LTP on 2015-16 CELDT

	2015-2016 CELDT Scaled Scores					
	Unmatched Results					*Matched Results*
Fixed Effects	Model 1 (Coef. *SE*)	Model 2 (Coef. *SE*)	Model 3 (Coef. *SE*)	Model 4 (Coef. *SE*)	Model 5 (Coef. *SE*)	Model 6 (Coef. *SE*)
Intercept/Constant	β=497.06***	β=531.53***	β=547.86***	β=398.75***	β=215.93***	β=310.59***
	(2.31)	(4.35)	(9.52)	(22.72)	(21.29)	(19.23)
Treatment						
Science LTP	β = 50.50***	β = 38.56**	β = 27.98~	β = 28.69~	β = 12.52	β =17.98***
(Non-Science LTP omitted)	(14.97)	(14.08)	(15.83)	(15.53)	(8.81)	(2.47)
Demographic Background						
Low Socioeconomic Status		β = -21.09***	β = -17.94***	β = -18.65***	β = -6.03*	β = -10.97***
(Non-low SES omitted)	-	(4.52)	(4.75)	(4.69)	(2.79)	(2.76)
ELL Designation						
(Non-ELL omitted)	-	-	-	-	-	-
SPED Classification		β = -28.47***	β = -25.79***	β = -28.37***	β = -23.53***	β = -25.34***
(Non-LTP omitted)	-	(4.87)	(4.92)	(4.85)	(2.81)	(2.50)
Female		β = 7.89~	β = 7.35~	β = 6.43	β = 3.53	β = 6.09**
(Male omitted)	-	(4.29)	(4.31)	(4.24)	(2.49)	(2.33)
Asian			β = -13.58	β = -16.70~	β = -16.92**	β = -11.57
(White omitted)	-	-	(9.86)	(9.83)	(5.85)	(5.27)
Black			β = -61.64**	β = -66.97	β = -20.29	β = -11.69
(White omitted)	-	-	(21.93)	(21.59)	(12.93)	(15.97)
Hispanic			β = -20.88*	β = -22.53**	β = -21.42***	β = -15.42**
(White omitted)	-	-	(9.47)	(9.44)	(5.60)	(4.10)
Native			β = 39.54	β = 36.01	β = 1.33	β = -34.02*
(White omitted)	-	-	(44.17)	(43.48)	(24.17)	(14.12)
Multi-racial			β = -25.60	β = -24.98	β = -31.47	β = -26.36
(White omitted)	-	-	(53.64)	(52.66)	(29.24)	(37.40)
Race-ethnicity Missing			β = -73.10***	β = -78.46***	β = -20.81~	β =-14.08
(White omitted)	-	-	(21.25)	(20.95)	(12.48)	(15.38)
2015-2016 Attendance				β = 0.88***	β = 0.01	β = -0.60***
(Continuous, Mean: 172)	-	-	-	(0.12)	(0.11)	(0.10)
Grade 5 CELDT					β = 0.66***	β = 0.67***
(Continuous, Mean: 505)	=	-	-	-	(0.02)	(0.02)
R^2	0.007	0.050	0.061	0.095	0.626	0.651

Key: ~ = p. < .10, * = p. < .05, ** p. < .01, *** = p. < .001

Note: These models included 57 LTP students and 1745 non-LTP students