Testing the Consensus Model of Effective PD: Analysis of Practice and the PD Research Terrain

Kathy Roth
Center for Excellence in Math and Science Teaching, California State Polytechnic University, Pomona, CA

Chris Wilson
Biological Sciences Curriculum Study, Colorado Springs, CO

Joe Taylor
Abt Associates, Cambridge, MA

Connie Hvidsten
Biological Sciences Curriculum Study, Colorado Springs, CO

Betty Stennett
Biological Sciences Curriculum Study, Colorado Springs, CO

Nicole Wickler
Center for Excellence in Math and Science Teaching, California State Polytechnic University, Pomona, CA

Susan Kowalski
Biological Sciences Curriculum Study, Colorado Springs, CO

Paul Numedahl
Biological Sciences Curriculum Study, Colorado Springs, CO

Jody Bintz
Biological Sciences Curriculum Study, Colorado Springs, CO
Rationale and Overview

The *Next Generation Science Standards* (NGSS Lead States, 2013) put extra pressure on our need for knowledge about models of professional development (PD) that support high-quality science teaching and learning. Traditional PD seldom provides teachers with the science content, pedagogical content knowledge (PCK), and pedagogical skills to help them teach in ways that support student development of scientific understandings and reasoning practices called for in the NGSS. Regarding PD for science teachers, in particular, the *Framework for K-12 Science Education* (NRC, 2012) emphasizes that successful implementation of the NGSS requires that teachers develop strong understandings of a) “the scientific ideas and practices they are expected to teach,” b) “the initial ideas students bring to school and how they may best develop understanding”, and c) “science-specific PCK” (p. 256).

Unfortunately, there is little research that examines the impact of various PD approaches on student learning. Absent this knowledge, researchers have proposed a consensus model of effective PD that includes the features described in figure 1 (Desimone, 2009; Wilson, 2013).

![Figure 1. Consensus Model of Effective Professional Development](image)

Few studies examine this consensus model empirically; fewer still look at impact on student learning (Little, 2011; Desimone, 2009; Yoon, Duncan, Lee, Scarloss, & Shapley, 2007). Recent reviews of PD sound strikingly similar to those of reports from a decade ago: There is consensus about what high-quality PD should look like and little empirical evidence to support that consensus model (Garet, Porter, Desimone, Birman, & Yoon, 2001; Wilson, 2013). But rigorous research that looks at impact on student learning is beginning to emerge, and the results do not always support the consensus model; programs that incorporate all of these features are not consistently demonstrating the desired impact on student learning in mathematics and reading (Garet et al., 2008; Garet et al., 2011; Hill, Beisiegel, & Jacob, 2013 Santagata, Kersting, Givvin, & Stigler, 2011). The paucity of research that looks at student learning outcomes, especially in science, limits our ability to support teachers in aligning their teaching practice with the NGSS (Reiser, 2013). The lack of consistent empirical data to support the consensus model of effective PD underscores the need for research that more clearly identifies key features of effective PD that are linked to student learning.

Another shortcoming of research on professional development is that we do not have lines of research that explore the effectiveness of PD programs for science teachers in increasingly rigorous ways over time. Borko (2004) mapped the terrain of research on PD in three phases. Phase 1 research involves the development of a PD program and initial evidence that it can positively influence teacher learning. In Phase 2, “researchers determine whether the PD program can be enacted with integrity in different settings and by different professional development providers” (p. 9). Phase 3 involves comparing the promising PD program to other PD programs, studying the impact on student learning as well as teacher learning, and identifying the resources needed for successful implementation across sites.
This paper reports on a rich line of research related to the Science Teachers Learning from Lesson Analysis (STeLLA) professional development program that explores the pathway of program influence on teacher science content knowledge, teacher PCK, teaching practice, and student learning in a series of quasi-experimental and experimental studies. The line of research explores aspects of all three phases of Borko’s research terrain but follows a slightly different path and includes an additional phase which examines the impact on student learning when key STeLLA program features are implemented in contexts that test the scalability and sustainability of the PD approach.

The paper maps the history of this line of research, demonstrating how it rigorously tests a videocase-based, analysis-of-practice professional development program, explores some of the territory in the three phases of research mapped out by Borko, and extends into a fourth phase that we describe as exploring scalability and sustainability. We also share insights about the consensus model of effective professional development revealed by this line of research. The paper is organized into four main sections:

Part 1 describes the initial development and study of the STeLLA PD program. It describes the STeLLA PD curriculum and structure and the results of an initial quasi-experimental study. The study fits in Borko’s Phase 1 as a development project that also produces evidence of effectiveness. It also moves into Borko’s Phase 3 by studying the impact on student learning (not just teacher learning) and by comparing the program to a content deepening PD program in a quasi-experimental study.

Part 2 describes the second study in this line of research. This randomized-controlled trial (RCT) study of the STeLLA program examined the same pathway of influence on teacher content knowledge, teacher PCK, teaching practice, and student learning, but the program was now led by professional development providers who were not part of the Phase 1 development work. In addition, the study was conducted in new geographic areas with a larger number of teachers (140) who were randomly assigned to treatment groups. Thus, this study moved into Borko’s Phase 2 territory, but it also pushed into Phase 3 as it included a comparison of the STeLLA Lesson Analysis program with a rigorous, content deepening PD program of equal duration, and it examined impact on student learning as well as teacher learning.

Part 3 describes how research related to the STeLLA PD approach extends beyond Borko’s three phases of PD research by exploring issues of scalability and sustainability. This section describes how three ongoing quasi-experimental studies are each adapting the STeLLA program to meet the needs of a new target audience, to enhance the potential for reaching more teachers/students, and to contribute to sustainability. The three contexts are (1) a program for preservice and first-year elementary science teachers, (2) a course for high school science teachers, and (3) a teacher leadership development program to support district-wide implementation of the STeLLA program in grades K–6 in a high-needs, urban district.

In Part 4 we examine ways in which the STeLLA PD did and did not incorporate the features of effective PD in the consensus model. We then identify key features of the STeLLA PD model that are not included in the existing literature descriptions of the consensus model. This analysis is then used to propose some features of PD that are worthy of attention in future research that we assert should use evidence of impact on student learning to produce a new, research-tested model of effective PD.

The STeLLA work represents an unusually rigorous line of research that has much to say about the kind of professional development needed to meet the goals for science learning defined in the NGSS. The STeLLA-related studies provide strong evidence that science teacher professional development can make a significant difference in terms of student learning, and they reveal key features of PD that push beyond the consensus model toward a research-based model of teacher learning.
Part 1: Phase 1 Research—Initial Development and Testing of the STeLLA Videocase-based, Analysis-of-Practice Program

Elementary teachers have little training in science-specific pedagogy and even less training in the scientific disciplines they are expected to teach (Dorph, Shields, Tiffany-Morales, Hartry, & McCaffrey, 2011; Fulp, 2002). In addition, elementary teachers consistently report concerns about their weak science content knowledge and their discomfort in teaching science (Wier, 1987). Yet these teachers now face new, higher standards as the Next Generation Science Standards (NGSS Lead States, 2013) challenge them to engage students in using science practices in more sophisticated ways to develop rich understandings of core science ideas and crosscutting concepts. How can elementary teachers develop the content knowledge, the PCK, and the teaching practices needed to meet this challenge?

To address elementary science teachers’ science learning needs, the Science Teachers Learning from Lesson Analysis PD program was designed, developed, and then tested in a quasi-experimental study. The program focuses on deepening upper elementary teachers’ content knowledge and pedagogical content knowledge in an integrated, practice-based way through scaffolded analysis of videocases of elementary science teaching and learning. The video analysis work is guided by a conceptual framework that supports teachers in learning to look at science teaching and learning through two lenses: the Student Thinking Lens and the Science Content Storyline Lens (figure 2). For each lens, teachers examine, analyze, and practice using a limited set of science teaching strategies that are drawn from literature on effective science teaching (seven Student Thinking Lens strategies and nine Science Content Storyline strategies). The videocases focus on specific science learning goals in the teachers’ curriculum, and analysis work is organized around a sequence of tasks that engage teachers in thinking about this science content in the context of learning how to reveal, support, and challenge student thinking and in the context of developing strategies for creating a coherent science content storyline within and across lessons.

An important guiding principle of this approach is that science content knowledge and pedagogical content knowledge will be more meaningful to teachers and have more impact on their teaching practice if they are embedded in analysis of practice. The use of videocases enables teachers to slow down the teaching and learning process and makes teaching practices more available for collaborative and rigorous inquiry and analysis.
STeLLA Conceptual Framework
Learning to analyze science teaching through two lenses

allows you to learn and use strategies for more effective science teaching

**SCIENCE TEACHING**

**STRATEGIES TO REVEAL, SUPPORT, AND CHALLENGE STUDENT THINKING**

1. Ask questions to elicit student ideas and predictions
2. Ask questions to probe student ideas and predictions
3. Ask questions to challenge student thinking
4. Engage students in interpreting and reasoning about data and observations
5. Engage students in using and applying new science ideas in a variety of ways and contexts
6. Engage students in making connections by synthesizing and summarizing key science ideas
7. Engage students in communicating in scientific ways

**STRATEGIES TO CREATE A COHERENT SCIENCE CONTENT STORYLINE**

A. Identify one main learning goal
B. Set the purpose with a focus question and/or goal statement
C. Select activities that are matched to the learning goal
D. Select content representations matched to the learning goal and engage students in their use
E. Sequence key science ideas and activities appropriately
F. Make explicit links between science ideas and activities
G. Link science ideas to other science ideas
H. Highlight key science ideas to other science ideas
I. Summarize key science ideas

Figure 2. STeLLA Conceptual Framework: The Two Lenses

**Development and Description of the STeLLA PD Program**

The program was designed to include features that had been called for in the literature around the turn of the millennium, when the STeLLA program was initially designed (figure 3) (Ball & Cohen, 1999; Carpenter, Fennema, Peterson, Chiang & Loe, 1989; Cobb et al., 1991; Cohen & Barnes, 1993; Cohen & Hill, 1998; Darling-Hammond & Sykes, 1999; Elmore, 2002 Garet et al., 2001; Kennedy, 1998; Lewis & Tsuchida, 1997, 1998; Loucks-Horsley, Hewson, Love, & Stiles, 1998; Mumme & Seago, 2002; Shimahara, 1998; Stigler & Hiebert, 1999; Takemura & Shimizu, 1993; Whitehurst, 2002; Wilson & Berne, 1999; Yoshida, 1999).

- Engages teachers actively in collaborative, long-term problem-based inquiries
- Treats content as central and intertwined with pedagogical issues
- Enables teachers to see these issues as embedded in real classroom contexts
- Focuses on the content and curriculum they will be teaching
- Is guided by an articulated model of teacher learning that specifies what knowledge and skills teachers will gain, what activities will lead to this learning, and how this new knowledge and skills will appear in their teaching practices

Figure 3. Design Features of the STeLLA PD Program
Although analysis of practice using video and other artifacts represents a form of PD that appears to have great potential, the research base provides limited guidance about the substance, or content, of effective analysis-of-practice programs. What should teachers be learning as they engage in analysis-of-practice activities? Kennedy (1998) found that studies of math and science teacher PD tend to focus on the form of the program, giving insufficient attention to program content, or substance. Form refers to how teachers’ PD activities are organized and how long they last (e.g., analysis-of-practice, one-year program, use of videocases, face-to-face or online). Substance, in contrast, refers to the ideas and skills that teachers learn about (e.g., knowledge about specific school subject matter, knowledge about how students learn specific subject matter, inquiry science teaching). In our description of the STeLLA program, we focus first on the program substance and learning goals and then describe the theory of teacher learning that guided decisions about how to integrate this program substance into a program form.

**STeLLA Program Substance.** The STeLLA program developers drew from research to select a small set of ideas that showed high potential for improving teacher learning, teaching practice, and student learning within a one-year program. To impact teaching and learning significantly within this one-year period, it was assumed that the substance of the program would need to be tightly focused on a few ideas addressed in depth. Given the multiple aspects of knowledge, skills, and dispositions that elementary science teachers need, the literature review produced a long list of candidate ideas for the STeLLA program content. The following five types of knowledge and abilities represent the learning goals for teachers in the STeLLA Lesson Analysis program. These learning goals are intertwined within the STeLLA conceptual framework (figure 2).

1. **Science Content Knowledge.** At each grade level, teachers are supported in deepening their understandings of science about a set of core science ideas in two topic areas. In the STeLLA program, work on developing these science content understandings occurs in the context of analyzing lesson videos and student work as well as in content deepening activities led by university science faculty. Teachers are also supported by a PCK-oriented content background document that describes science content ideas in relevant teaching contexts such as how common student ideas are related to science ideas, how commonly used lesson activities develop or fail to develop core science ideas, or how content representations and analogies have strengths and weaknesses in terms of scientific accuracy and in terms of supporting student learning.

2. **Pedagogical Content Knowledge.** The STeLLA development team drew from a broad literature review to select a focus on two types of pedagogical content knowledge that were hypothesized to have the most potential to impact both teacher and student learning within a one-year time frame: 1) knowledge about creating a coherent science content storyline and 2) knowledge about eliciting, supporting, and challenging students’ thinking about specific science content.

   a. **Science Content Storyline Lens.** This lens focuses attention on how the science ideas in a science lesson or unit are sequenced and linked to one another and to lesson activities to help students construct a coherent “story” that makes sense to them. Many times, teachers present accurate science content and engage students in hands-on activities, but these ideas and activities are not woven together to tell or reveal a coherent story. The TIMSS Video Study found that eighth-grade US science lessons focus on doing activities with less attention to the science content and even less attention to the links between activities and science ideas. When compared to lessons in higher-achieving countries, US lessons present poorly organized content that is weakly linked to activities, if at all (Roth et al., 2006). In higher-achieving countries, teachers more commonly use activities to develop science ideas and organize lessons in a way that resembles a storyline—with clear and explicit connections made between the opening focus question, the science ideas, the activities, the follow-up discussions of activities, and the lesson ending. This contrast between US lessons and those in higher-achieving countries highlights the need for teachers to develop
pedagogical content knowledge about ways to organize specific science content and activities so that students can see and make the links between science ideas and lesson activities. The research team hypothesized that the reform call for more inquiry-oriented science teaching was being interpreted by teachers as a call for doing activities, while an essential aspect of scientific inquiry—the development of science ideas—was being neglected. The STeLLA Science Content Storyline Lens was brought to life for teachers in the STeLLA program through a set of nine teaching strategies to support the development of a coherent storyline (see figure 2). These strategies are described in a booklet that teachers use as a reference throughout the program.

b. Student Thinking Lens. The decision to develop teachers' pedagogical content knowledge about students' ideas and ways of thinking about specific science content standards was influenced by the extensive body of research that explores 1) students' ideas about natural phenomena, 2) the role student ideas can play in teachers' science teaching, and 3) the importance of attending to student ideas and cultures in supporting science learning of students from groups underrepresented in science. Because teachers do not typically focus on students' ideas in their planning, teaching, and analyses of teaching (Sherin & van Es, 2002), they need to develop the disposition and ability to pay attention to student thinking; to do this effectively, they also need knowledge of common student ways of thinking about scientific phenomena. This analytical disposition and pedagogical content knowledge can enable teachers to anticipate student ideas and difficulties and respond to them in ways that will help students move forward in their understanding of important scientific ideas and practices. Seven teaching strategies that reveal, support, and challenge student thinking are described in the STeLLA strategies booklet.

3. Ability to Use Content and Pedagogical Content Knowledge to Analyze Teaching and Learning. To support teachers in becoming analytical about their science planning and teaching, the STeLLA curriculum engages teachers in learning how to analyze practice using a lesson analysis process, lesson analysis norms, and a set of analytical tools.

a. STeLLA Lesson Analysis Process. Paralleling the kinds of reasoning used in developing scientific explanations, the STeLLA lesson analysis process begins with turning an observation of a video clip or of student work into a claim about something related to either the Student Thinking Lens or the Science Content Storyline Lens. Next, the teacher supports the claim with specific evidence from the video (e.g., pointing to a line or section in the transcript) or from student written work. The teacher then provides reasoning to support the claim by tying the evidence to scientific principles, to the STeLLA strategy booklet, to research on science teaching, to reform documents such as standards, to their personal experience, or to other resources. Finally, the teacher proposes an alternative in the form of either an alternative claim or an alternative teaching move. Figure 4 is a visual depiction of this lesson analysis process.
b. **STeLLA Norms and Analysis Basics.** To support the development of a learning community within the study groups where teachers can feel safe to have their ideas and teaching analyzed and challenged, STeLLA teachers discuss and agree upon a set of norms that includes both basics (arrive prepared and on time and stay for the duration; eliminate interruptions such as cell phones and email; make room for participation from all—monitor floor time; remain attentive, thoughtful, and respectful) and norms that are at the heart of the STeLLA program (keep the goal in mind—it’s all about improving student learning; share your ideas, uncertainties, confusions, disagreements, and questions; expect and ask challenging questions to deepen everyone’s learning). In addition, each teacher study group is encouraged to develop their own explicit norms to create a productive and safe learning environment.

In addition to these general norms, teachers also share a set of norms specific to video analysis work:

**Viewing Basics**
- Look past the trivial, the little things that “bug” you.
- Avoid the “this doesn’t look like my classroom” trap.
- Avoid making snap judgments about the teaching or learning in the classroom you are viewing.

**Analysis Basics**
- Focus on student thinking and the science content storyline.
- Look for evidence to support any claims.
- Look more than once.
- Consider alternative explanations and teaching strategies.

c. **STeLLA Lesson Analysis Tools**

i. **Analysis Guides.** When teachers are first introduced to the Science Content Storyline Lens teaching strategies, analysis guides scaffold their efforts to dig deeply into the meaning of each of these strategies. Each strategy has its own analysis guide. For example, the analysis guide for Strategy D (*select content representations matched to the learning goal and engage students in their use*) prompts teachers to consider whether a content representation is...
scientifically accurate, is closely matched to the learning goal, presents ideas in ways that are comprehensible to students, addresses or introduces/reinforces common student misconceptions, and/or distracts students from the main learning goal with too many details or new terms. To analyze student use of the representations, the guide asks teachers to consider whether students are involved in modifying, creating, analyzing, or critiquing the representation.

ii. **Lesson Analysis Protocol.** When teachers begin to analyze video clips and student work from each other’s classrooms, they use a Lesson Analysis Protocol to keep the discussions focused and productive (figure 5). After first identifying the use of particular STeLLA strategies in the video (or other videocase artifact), teachers individually respond to an analysis focus question by writing down a claim, evidence, reasoning, and an alternative (CERA). The focus question is often focused on student thinking: *What student thinking was made visible through the use of the identified strategy?* The teachers’ individual writing provides rich fodder for the discussion to follow. The process ends with the videotaped teacher reflecting on what she or he learned and noticed from listening to the group discussion.

### Lesson Analysis Protocol

<table>
<thead>
<tr>
<th>1. Identify the Purpose and Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>- What instances do you note of Strategies 1, 2, or 3? (Ask questions to elicit student ideas and predictions, Ask questions to probe student ideas and predictions, Ask questions to challenge student thinking)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Analysis Focus Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>- What did the use of STeLLA questioning strategies reveal about students’ understanding or misunderstanding about the lesson focus question?</td>
</tr>
<tr>
<td>- In particular, what student thinking was made visible because of the use of challenge questions? What were missed opportunities?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lesson Analysis Step</th>
<th>To Do</th>
<th>Your Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Claim</td>
<td>Turn an observation, question or judgment into a specific claim that responds to the focus question.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Evidence and Reasoning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point to a specific place in the video transcript, lesson plan, or student work that supports your claim.</td>
</tr>
<tr>
<td>Connect your claim and evidence with reasoning based on research or wisdom of practice. Also look for evidence that challenges your claim.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Alternatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consider an alternative interpretation or explanation.</td>
</tr>
<tr>
<td>Consider new questions this might raise.</td>
</tr>
<tr>
<td>Consider alternative question(s), activity(s), or strategies</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Reflect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facilitator and participants share reflections on the analysis discussion.</td>
</tr>
</tbody>
</table>

Figure 5. STeLLA Lesson Analysis Protocol

iii. **Features Analysis Charts.** STeLLA teachers regularly analyze student thinking in lesson videos and in student work. But they are also supported in using a more formal, systematic process to study student learning and progress based on students’ responses to open-ended
items on pretests and posttests. Using a Features Analysis Chart (figure 6), teachers first create ideal student responses to the open-ended item. They then deconstruct the ideal response into a list of specific aspects that might be present in an ideal answer. In addition, they list common student confusions and misunderstandings that might appear in student responses. The Features Analysis Chart is set up so that the teacher can then check off which ideas/confusions are held by each student on the pretest and on the posttest. The Features Analysis Chart provides a good visual representation of progress from pretest to posttest both by the class as a whole and for individual students. This process keeps the focus off of looking just for one right answer to a new focus on tracing student thinking, looking for development and progress, and identifying specific areas that need additional work.

![Features Analysis Chart](image)

Figure 6. STeLLA Assessment Features Analysis Chart

4. **Ability to Use Content and Pedagogical Content Knowledge in Teaching Science.** We assume that content knowledge and pedagogical content knowledge about science content storylines and about student thinking will only impact student learning if teachers are able to apply that knowledge in their teaching. To support teachers’ initial efforts at using the STeLLA lenses and teaching strategies in their own classrooms, teachers are given a set of six lessons for use in their fall teaching about either Earth’s changing surface (4th grade) or Sun’s effect on climate/seasons (5th grade). The lesson plans, which teachers first preview and analyze in the summer institute, are formatted in ways that highlight the STeLLA lenses and strategies and that clarify key issues related to the science content. For example, there is explicit attention throughout the lesson plan to anticipated student responses and to follow-up probe and challenge questions the teacher might ask. These lesson plans were designed by the research team to be educative curriculum materials that model and scaffold teachers’ initial use of
the STeLLA strategies and support teachers in deepening their science content knowledge (Davis & Kracjik, 2005). These lesson plans serve three important functions: 1) They provide modeling and scaffolding for teachers in using the two lenses and the supporting teaching strategies, 2) they provide more opportunities for teachers to clarify their science content understandings as they teach, and 3) they provide a common curriculum for the lesson analysis work since the teachers in the study were using different science curriculum materials in their schools/districts.

During the last semester of the program (winter/spring of the school year), teachers are not provided with STeLLA lesson plans. Instead, they work collaboratively to develop lesson plans in a second content area, using their learning about the science content and about the STeLLA lenses and strategies. Teachers are supported in this planning work through content deepening activities and tools (e.g., analysis of lesson videos in this content area, content background document written in a PCK-focused style) and by STeLLA planning tools that keep them focused on using the STeLLA strategies to plan coherent, learner-focused lessons and to guide their teaching of the science lessons.

**STeLLA Program Theory of Teacher Learning.** Decisions about how to design the program form in order to meet these teacher learning goals were guided by a situated cognition model of teacher learning, together with a cognitive apprenticeship model of instruction. Situated cognition posits that learning is naturally tied to authentic activity, context, and culture (Collins, 2006; Lave & Wenger, 1999). In line with the cognitive apprenticeship model, the STeLLA program is organized to support classroom-based (“situated”) learning over time, with more direct scaffolding by PD leaders at the beginning of the program that moves gradually toward more teacher-directed work at the end of the one-year program. The following description of the STeLLA program form illustrates how the design of the program was driven by this theory of teacher learning to scaffold teachers’ learning of the program substance over time through a sequence of analysis-of-practice, planning, and teaching activities.

**STeLLA Program Form.** Central to the form of the program was teachers’ interactions with each other and with videocases as they analyzed science content, science teaching, and science learning in small (5–10 teachers), grade-level study groups, each led by a PD leader. These interactions began during a two-week summer institute and continued in monthly 3.5-hour meetings across the school year.

During the summer institute, lesson analysis work was scaffolded by a carefully planned sequence of tasks that engaged teachers in learning about and using the STeLLA lenses and strategies, the science content and the lessons they would be teaching in the fall, and the lesson analysis process. In this work, teachers analyzed videocases of teachers and students outside their own study groups. Summer institute work also included time spent with university science faculty who engaged teachers in a variety of inquiry, model-building, and problem-solving activities to develop understandings of the target science ideas in the two content areas that teachers would be teaching during the school year. Thus, teachers had opportunities to deepen their content knowledge in two contexts—lesson analysis activities and activities focused on content deepening.

During the fall teachers taught the STeLLA lesson plans and collected student work from this teaching. In three monthly, 3.5-hour study group meetings they learned from analysis of videos from each other’s classrooms and from analysis of their students’ work. The scaffolding from the PD leader remained high. The PD leader selected video clips and focus questions for analysis, used the Lesson Analysis Protocol to guide, support, and challenge teachers’ analyses, and made decisions about when and how to address science content confusions that arose during the lesson analysis process.

During the winter and spring, each study group switched to a new content topic area. The PD leaders’ scaffolding and structuring remained but was gradually reduced. For example, the teachers were not given STeLLA lesson plans for these new content areas. Instead, the teachers worked in the study groups and on
their own outside of the study group to plan a sequence of connected lessons in this second content area. STeLLA planning tools and strategies kept the focus on the STeLLA framework and supported teachers with the challenging task of creating coherent science content storylines within and across lessons. Videocases from other teachers’ classrooms were used to continue to deepen teachers’ understanding of this content area and of the STeLLA strategies.

Table 1 presents an overview of the form and substance of the STeLLA professional development program.

Table 1. Form and Substance of the STeLLA Professional Development Program

<table>
<thead>
<tr>
<th>Program structure</th>
<th>Program form</th>
<th>Program substance</th>
</tr>
</thead>
</table>
| **Summer institute 2 weeks (9 days)** | Half-day content deepening sessions (3.5 hrs/day)  
- Grade-level study groups  
- Led by university science faculty  
- Focus on two science content areas teachers will teach during the following school year  
- Activities: variety of content deepening activities | Science content knowledge in two topic areas in teachers’ curriculum |
| **58.5 hours face-to-face PD** | Half-day lesson analysis sessions (3 hrs/day)  
- Grade-level study groups  
- Led by BSCS STeLLA PD leaders  
- Focus on content area to be taught in the fall  
- Activities  
  - Analysis of video and student work from other teachers’ classrooms  
  - Analysis of lesson plans to be taught in the fall | Introduction to the STeLLA Student Thinking Lens and Science Content Storyline Lens and teaching strategies (PCK)  
Lesson analysis process, norms, and tools  
Science content deepening in the content area to be taught in the fall  
Lesson plans to be taught in the fall |
| **Academic-year sessions (monthly study group meetings)** | Three fall sessions (3.5 hrs each)  
- Grade-level study groups  
- Led by BSCS STeLLA PD leaders  
- Focus on content area taught in the fall  
- Activities  
  - Analysis of lesson video from participating teachers’ teaching of the STeLLA lesson plans  
  - Analysis of student work from participating teachers’ teaching of the STeLLA lesson plans | Deeper understanding of the STeLLA Student Thinking Lens and Science Content Storyline Lens and teaching strategies (PCK)  
Advancing lesson analysis and student work analysis abilities  
Science content deepening in the content area taught in the fall  
Ability to use STeLLA strategies in teaching practice |
| **28 hours of face-to-face PD** | Five winter/spring sessions (3.5 hrs each)  
- Grade-level study groups  
- Led by BSCS STeLLA PD leaders  
- Focus on content area taught in the spring  
- Activities  
  - Analysis of lesson video and | Deeper understanding of the STeLLA Student Thinking Lens and Science Content Storyline Lens and teaching strategies (PCK)  
Advancing lesson analysis and student work analysis abilities |
<table>
<thead>
<tr>
<th>Program structure</th>
<th>Program form</th>
<th>Program substance</th>
</tr>
</thead>
</table>
| Culminating session | student work from other teachers’ classrooms  
Content deepening activities  
Collaborative lesson planning activities | Science content deepening in the content area taught in the fall  
Ability to use STeLLA strategies in teaching practice |
| PD | All study groups meet together to share learning, video clips, analysis work; summarize key learning and next steps | Synthesis |
| Total PD hours | 88.5 hours | | |

Summary of Program Substance and Form. There are several constants throughout the STeLLA PD program that reflect the key principles that guided the design of the STeLLA program (figure 3). First, the STeLLA program has a strong storyline that focuses throughout on a conceptual framework that highlights the Student Thinking Lens and Science Content Storyline Lens and supporting teaching strategies (figure 2). Other constants are that teacher learning is situated in practice, focuses on the particular science content that teachers are teaching, and intertwines teacher learning about subject matter content with their learning about pedagogy as they analyze practice via lesson videos and student work.

Consistent with the situated cognition theory of teacher learning and the cognitive apprenticeship instructional model, what varies over time are the amount and type of scaffolding changes, the nature of the teachers’ activities, and the nature and role of the videocases. In the summer, teacher learning is highly scaffolded by the PD leaders, by the carefully selected videocases from other teachers’ classrooms, by the supporting resources and tools (such as the booklet describing each of the STeLLA strategies, the lesson analysis process, the analysis guides, the PCK content background document), and by a carefully planned set of lesson analysis and content deepening activities that are led by PD leaders and university science faculty. In the fall, the videos from the participating teachers take center stage so the content of the study groups becomes more driven by what happened in teachers’ classrooms. PD leaders are still heavily involved as they select the video clips and design the analysis focus questions, but they now have to be more flexible in the unfolding of the study group session, encouraging and allowing teachers to take over the lead in the analysis as much as possible but stepping in to clarify and challenge to support deep teacher learning. In the winter/spring work, teachers again look at videocases from other teachers’ classrooms, now in a new content area, but now they are focusing on using those analyses to deepen their content knowledge in this new content area and to generate ideas about how to plan lessons that will support students in developing strong understandings of the content. Scaffolding from STeLLA PD leaders and STeLLA tools remain, but teachers take on more leadership and ownership of the lesson analysis and planning process. Thus, both the program substance and the program form are carefully designed to support teacher learning and growth as effective science teachers.

Phase 1 Research: Existence Proof Study of STeLLA

The program was implemented and studied following a pathway of influence that included impact on teacher content knowledge, teacher pedagogical content knowledge, teaching practice, and student learning outcomes (figure 7). In a quasi-experimental design, 48 upper-elementary teachers in urban settings in Southern California volunteered for either the year-long STeLLA Lesson Analysis program or a Content Deepening program that focused on the same two content areas being addressed in the Lesson Analysis program. After participating in separate but equally intensive experiences in a two-week summer institute, teachers in both groups addressed the same set of learning goals as they taught the two content areas to their students during the school year. The Lesson Analysis teachers met in study groups across
the school year, studying video clips and student work from their teaching of these two content areas. The Content Deepening teachers received no further science professional development. At three points in time (pre-, mid-, and post-program) teachers in both groups took a content knowledge test about the focus content and demonstrated their pedagogical content knowledge by analyzing video clips of science teaching about the focus content. To assess changes in teachers’ practice, lesson videotapes in STeLLA teachers’ classrooms were collected before and after their participation in the program. Student learning was assessed by pre-post unit tests given the year before the STeLLA program and again on the same content at the end of the program.

Findings/Analysis. Research findings (Roth et al., 2011) showed that, in comparison with teachers who received science content deepening PD only (n = 16), teachers experiencing the one-year STeLLA PD lesson analysis program (n = 32) developed deeper science content knowledge (p < .001) and stronger abilities to use PCK to analyze science-teaching practice (p < .001). In addition, teachers in the STeLLA program increased their use of teaching strategies that made visible student thinking and contributed to the coherence of the science lesson (p < .01). Most importantly, their students’ learning showed significant improvement (p < .01, average effect size d = 0.47). Hierarchical Linear Modeling (HLM) analyses revealed that teachers’ science content knowledge, ability to analyze student thinking, and use of four science content storyline teaching strategies were predictors of student learning: 1) identify one main learning goal, 2) select content representations matched to the learning goal and engage students in their use, 3) make explicit links between science ideas and activities, and 4) link science ideas to other science ideas.

The study found that a one-year PD program can result in significantly improving students’ science learning. In addition, the study identifies the importance of teachers’ content knowledge, their ability to analyze student thinking, and their use of four particular teaching strategies related to the Science Content Storyline Lens. We propose that these four strategies are good nominations as high-leverage science teaching practices (Windschitl, Thompson, Braaten, & Stroupe, 2012). This study represents the beginning of an all-too-rare line of research that measures the impact of PD in terms of entire pathway of influence, including impact on student learning.
Part 2: Phase 2 and 3 Research—An RCT Study Comparing Analysis-of-Practice PD and Content Deepening PD

In the original study of the STeLLA professional development program (Part 1), student effects were estimated using a quasi-experimental design, the program was delivered directly by the program developers, and the STeLLA lesson analysis PD was compared to a content deepening program that was of shorter duration. The current study addresses these design limitations by testing the efficacy of STeLLA using a randomized design, increasing the internal validity of the study conclusions, and exploring whether the STeLLA program is as effective when delivered by those outside the original development team, when compared with a strong content deepening PD program of equal duration, and when delivered to a larger number of teachers who are randomly assigned to the STeLLA program or the content deepening program.

Design/procedure. Using the same teacher and student outcome measures as the original STeLLA study (figure 4), this study of the STeLLA PD program employed a cluster-randomized design (Raudenbush, 1997) where schools were the units (clusters) assigned randomly to either the Lesson Analysis or Content Deepening intervention groups. In this paper we report the student learning outcomes. The student content knowledge outcome measures were converted to Rasch person measures and compared across intervention groups. Doing so puts the scores on an equal interval scale that is most appropriate for use in inferential statistics. We quantified the STeLLA intervention effect by comparing across intervention groups the pretest-adjusted posttest scores from these content knowledge measures. The pretest adjustment is to account for any pre-intervention differences across treatment groups. To account for the nested nature of the data (i.e., students and teachers within schools), we used multilevel modeling to make the adjusted posttest mean comparisons (Raudenbush & Bryk, 2002).

In addition to testing the statistical significance (α = 0.05) of the STeLLA intervention effects, we estimated the practical significance by computing a Hedges’ g effect size appropriate for effects estimated from multilevel analyses.

Findings. Table 2 includes the level-two parameter estimates from that analysis, including the covariate-adjusted treatment effect for students.

Table 2. Level Two Parameter Estimates: Students (Fixed Effects)

<table>
<thead>
<tr>
<th>Level 2 fixed effects</th>
<th>Unstandardized coefficient</th>
<th>Standard error</th>
<th>t-ratio</th>
<th>DoF</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTERCEPT</td>
<td>53.22</td>
<td>0.42</td>
<td>127.77</td>
<td>74</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td>TREAT</td>
<td>6.11</td>
<td>0.84</td>
<td>7.27</td>
<td>74</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td>PRE</td>
<td>0.39</td>
<td>0.15</td>
<td>2.66</td>
<td>74</td>
<td>p = .01</td>
</tr>
</tbody>
</table>

The unstandardized coefficient for treatment is the pretest-adjusted estimate of the difference in mean achievement between the two treatment groups, and that value is about 6% (6.11) in favor of the STeLLA Lesson Analysis treatment group students. This effect is statistically significant at p < .001 with a 95% confidence interval of [4.46% < --- > 7.76%]. This corresponds to a Hedges’ g effect size of 0.68 standard deviations with effect size confidence interval [0.60 < --- > 0.76].
This effect is consistent with those observed in the quasi-experimental comparisons made before the test equating process. For example, the two quasi-experimental subsample comparisons that had sufficient statistical power (i.e., involved a sufficient number of schools) to have stable standard errors and thus make defensible claims were those that used the Food Webs and Water Cycle outcome assessments. For the Food Webs comparison, the treatment effect estimated from equations 1 and 2 was \( \beta_{01} = 2.47 \) (\( p < .001 \)) and the effect size was \( g = 0.58 \). For the Water Cycle comparison, the corresponding effects were \( \beta_{01} = 3.11 \) (\( p < .001 \)), \( g = 0.75 \). In the following sections, we extend the discussion of the confirmatory main effect of treatment (\( g = 0.68 \)) by situating it in the larger context of elementary intervention effects and by describing other metrics of its practical significance.

Student effect sizes such as these compare well to empirical benchmarks for effect sizes. For example, Hill, Bloom, Black, and Lipsey (2008) conducted a synthesis of effect sizes for randomized control trials and found that for elementary school studies the average effect size was 0.33. Further, they found that the average effect size varied by the breadth of focus for the outcome measure, reporting on page 8: “within studies of elementary schools, mean effect sizes are highest for specialized tests (0.44), next-highest for narrowly focused standardized tests (0.23), and lowest for broadly focused standardized tests (0.07).” Given that the student effect size reported in this study (0.68) was computed using scores from a narrowly focused standardized test, but one justly focused on the project learning goals (i.e., equally accessible to both treatment groups), we find that their magnitude compares favorably to those of similar studies.

Another way to interpret these effect sizes is to compare them to normative expectations for achievement growth (i.e., average pre-post year effect sizes for elementary school science students). This effect size expresses students’ expected gain in science achievement in the absence of an intervention over and above business-as-usual (BaU) practices. Looking across a set of nationally normed tests, Hill et al. (2008) estimated that the average pre-post year effect size for science in the grades 4–5 is 0.40 standard deviations. The effect size of 0.68 detected in this study is noteworthy as it corresponds to 0.68/0.40 (170%) of the one-year expected gain for these students. Thus, we estimate that Lesson Analysis students emerge from the study nearly two years ahead of the comparison group students in achievement on these science topics.

As a final way to express the practical importance of the treatment effect, we converted the effect size into an improvement index using the properties of the normal distribution. In a normal distribution, a 1.0 standard deviation effect size is equivalent to 34 percentile points. Therefore, an effect size of 0.68 equates to an improvement index of 23 (34 \( \times \) 0.68) percentile points. So, if the Content Deepening students were at the mean of a normed sample, the 50th percentile, the Lesson Analysis students would then be placed at the 73rd percentile.

Despite well-documented difficulties in detecting the effects of professional development on student outcomes, this study yielded strong evidence that the STeLLA Lesson Analysis approach can be efficacious for students, even when nondevelopers deliver the intervention. We are currently examining the effects of STeLLA on teacher outcomes (content knowledge, PCK, and teaching practice). Such analyses of teacher outcomes will enrich future study of student effects as these teacher outcomes can be tested for their influence as mediators of the STeLLA effect on students and for their ability to identify particular features of the professional development and particular teaching strategies that predict student learning.
Because study of the STeLLA program demonstrated important impact on students’ science learning, the next goal in this line of research is to investigate issues of scalability and sustainability. As described in Parts 1 and 2, STeLLA is an intensive PD curriculum that has many components that need to be well aligned. How can such a program reach large numbers of teachers and students? This section describes three studies in which we are exploring the scalability and sustainability of the STeLLA PD model in three different ways.

Theoretical Perspectives about Scalability and Sustainability

STeLLA II was designed as a rigorous scale-up study of the kind prioritized in the last decade by education researchers and research funders (Feuer, Towne, & Shavelson, 2002; McDonald, Keesler, Kauffman, & Schneider, 2006; IES, 2012; US Dept. of Education, 2002). This study was a scale-up in the sense that it was designed to reach more teachers/students/schools/districts in a new geographic context and with new PD providers leading the program. It was also a scale-up in terms of its randomized, controlled research design. This scale-up study provided strong evidence of impact on both teacher and student science learning. Based on these results, we are now interested in reaching more teachers and students. But instead of simply replicating and studying STeLLA in additional RCT studies in different geographic contexts, we are now studying STeLLA’s impact in three quasi-experimental design studies to address two main research questions:

- Can the STeLLA approach be adapted to support the science learning of new target populations of teachers and students (high school teachers and students; preservice/first-year teachers and their students; and K–3 teachers and students)?
- Can the STeLLA approach be adapted in ways that will support more widespread dissemination and that will make it more likely to be sustainable?

In exploring these questions, we draw from three theoretical perspectives about scale-up in education (Coburn, 2003; McDonald et al., 2006; Lynch, 2012). McDonald et al. (2006) describe scale-up primarily in terms of research that collects valid and reliable data to show whether or not reforms can be implemented successfully in increasingly large and varied contexts. This was the perspective on scale-up that we tested in the STeLLA II RCT study, and we continue to be interested in reaching more teachers and students in a wider array of contexts.

But we are also interested in questions about scaling up suggested by the theoretical contributions from Coburn (2003) and Lynch (2012). Coburn conceptualizes scale-up as going beyond the focus on increased numbers to include four interrelated dimensions: depth of impact on practice; sustainability beyond the start-up funding period; spread that includes beliefs, norms, and principles to additional classrooms and schools; and shift in reform ownership from external partners to teachers, schools, and districts. Lynch uses activity theory to examine the overlap and the conflict or competition between the ways in which different players in the partnership (students, teachers, school districts, researchers) experience the goals of the reform effort and how they support the activities of the reform implementation. When there is agreement and overlapping support among these players related to the goals, tools, and activities of the reform effort, the intervention is poised for success. When two or more of these partnership systems are out of sync with each other or when any one system is faced with competing goals, tools, and activities, the success of the intervention is put at risk.
In this section we describe the three studies that are exploring issues of scale-up and sustainability of the STeLLA program. Drawing from the McDonald et al. (2006), Coburn (2003), and Lynch (2012) perspectives on scale-up, our project descriptions will consider the following questions:

- What is the context and target population for this study, and how is STeLLA adapted for this audience?
- What are the goals for depth of impact, and how are we studying depth of impact?
- How will this study help us learn about the potential for spread of the STeLLA PD approach, including the spread of ideas (such as the STeLLA conceptual framework and its teaching strategies)?
- How does this study explore/address issues of sustainability of the reform effort?
- Is there a plan for shifting ownership of the reform to the schools, districts, or university partners?
- To what extent do students, teachers, school districts, and researchers who are involved in the project share common goals and support activities related to the success of the reform effort?
- What are we learning about competing goals and activities across these groups that threaten the success of the reform effort?

**Study 1 Context: ViSTA Plus Program for Preservice and First-Year Teachers**

*a. What is the context and target population, and how is STeLLA adapted for this audience?*

The STeLLA PD program was designed for in-service elementary teachers. Can the program be adapted to have similar impact on the learning and professional growth of preservice and first-year teachers and their students? In the Videocases for Science Teaching Analysis Plus (ViSTA Plus) project, we took the STeLLA framework, tools, and resources and designed a variation of the STeLLA PD program for elementary preservice teachers who are students in university teacher education departments. The ViSTA Plus program includes a semester-long analysis-of-practice methods course that introduces preservice teachers to STeLLA strategies and lesson analysis and prepares them for participation in BSCS-led study groups that continue to meet synchronously online during student teaching and the first year of teaching.

The practice-focused science methods course is organized around the STeLLA conceptual framework and engages preservice teachers in learning from practice. In this sense, the course turns traditional teacher education upside down, immersing preservice teachers in practice from the beginning of the program. Teachers learn from practice in different ways in the methods course. First, they analyze videos and student work from other teachers’ classrooms to support their learning about science and particular science concepts, about the STeLLA Student Thinking and Science Content Storyline Lenses, and about the STeLLA teaching strategies. This work is similar to what in-service teachers experience in the STeLLA summer institute—analyzing videos and student work from other teachers’ classrooms. But the methods class also includes additional opportunities for preservice teachers to practice using the ideas they are learning from their analyses. For example, preservice teachers interview students about their science ideas to practice asking questions that probe student thinking; they analyze lesson plans that highlight the STeLLA strategies and practice teaching these lessons to each other; and they teach science lessons in elementary classrooms. During student teaching, they teach the STeLLA lesson plans in elementary classrooms and meet synchronously online to analyze and learn from video clips from each other’s lessons. This study group work continues during the first year of teaching, first using the same lesson plans they taught during student teaching and then venturing into a new content area. This study group work is similar to what STeLLA in-service teachers experienced in the year following their participation in a two-week summer institute.
b. What are the goals for depth of impact, and how are we studying depth of impact?

The goals for depth of impact are the same as those for in-service teachers in the STeLLA PD. We expect to deepen preservice teachers’ understanding of specific science concepts and scientific ways of knowing, to support the development of pedagogical content knowledge about students’ thinking and about the science content storyline, to support teachers in developing the skills and dispositions to learn how to learn from analysis of practice, and to enact teaching strategies that will enable their students to develop rich understandings. In short, we expect to produce well-started beginning teachers who are demonstrating greater impact on student learning than teachers who experience traditional teacher preparation experiences.

To assess our success in having these kinds of impact, we are studying the ViSTA Plus program in partnership with two universities and all of the school districts where ViSTA Plus teachers begin their teaching careers. We are comparing the experiences of preservice/first-year teachers in the ViSTA Plus program with a comparable group of teachers who experience the business-as-usual teacher education program at the same two universities (methods, student teaching) and who receive business-as-usual support during their first year of teaching within their respective school districts. This quasi-experimental study is using the STeLLA assessment tools to examine impacts of this preservice version of the STeLLA program on beginning teachers’ science content knowledge, on their ability to analyze video clips of science teaching using pedagogical content knowledge, on their teaching practice, and on their students’ learning. Videos of their science teaching during student teaching and during the first year of teaching will be used to describe their teaching practice.

c. How will this study help us learn about the potential for spread of the STeLLA PD approach, including the spread of ideas (such as the STeLLA conceptual framework and its teaching strategies)?

We need teachers who can be effective with students from the beginning of their careers. This study tests whether we can accomplish this goal by working within the traditional university structures of a methods course and student teaching. It will also help us learn about what might be gained when the curriculum and the learning community structures established in the university (in this case, the ViSTA Plus program and the BSCS-led study groups) are supported post-graduation and throughout the first year of teaching. This study will help us understand what is possible for beginning teachers to accomplish in their teaching when they are supported in learning from analysis of practice. It will also provide important insights about ways in which the STeLLA program needs to be modified to meet the needs of preservice/first-year teachers.

d. How does this study explore/address issues of sustainability of the reform effort?

e. Is there a plan for shifting ownership of the reform to the schools, districts, or university partners?

f. To what extent do students, teachers, school districts, and researchers who are involved in the project share common goals and support activities related to the success of the reform effort? What are we learning about competing goals and activities across these groups that threaten the success of the reform effort?

This study is a scale-up effort in the sense that it is exploring the adaptation of STeLLA PD to meet the needs of a new target audience—preservice/first-year teachers. But regarding sustainability and additional scale-up at more sites, this study raises questions and challenges rather than explores solutions. The project is currently driven and coordinated by BSCS and supported strongly by the methods course instructors who implement the ViSTA Plus methods course at the two universities. But the success of the project depends on the support and coordination of many people—at the universities, in the schools/districts that support the student teachers, and in schools/districts where these teachers begin their
teaching careers. There are definitely places where goals and activities of the program are in conflict or in competition with goals/needs of the universities and schools. A key example of this is the need for all the ViSTA Plus teachers to teach the same content to their students during student teaching and during the first year of teaching. This is critical in order to deepen teachers’ understanding of the content itself so that they can, in turn, engage in productive analysis of student thinking and learning and of the effectiveness of the science content storyline. But, not surprisingly, the selected content does not always match the content in the curriculum in participating schools. To successfully negotiate permission for the teaching of these lessons in these sites, we reduced the number of lessons to 3–5, recognizing that this may or may not be sufficient to support both the teachers’ and the students’ learning. Another example of competing goals is the content of the methods course. Methods course instructors are invested in their own approaches and curriculum, reflecting a tradition in which methods course instructors construct their own courses. For many methods course instructors, studying their own course is their primary research endeavor, so it is critical for their career advancement and tenure. In particular, methods course instructors are reluctant to drop things from their course syllabus and to adopt the ViSTA Plus philosophy of focusing in depth on analysis of practice.

Might the VISTA Plus approach ever be “owned” by universities and school districts? We would love to see universities embrace this curriculum, make study group development a central feature of their student teaching experiences, and see their mission as continuing to support these study groups into the first year of teaching. Our hope is that the results of our study—especially its impact on beginning teachers’ practice and elementary students’ learning—will open up conversations about reimagining university teacher education programs. We recognize, however, that this will be an uphill battle because it challenges so many well-ingrained practices of university teacher education, because it affects so many people, and because it has financial implications. It will also be interesting to see if the ViSTA Plus approach will be attractive for use in alternative pathways into teaching.

**Study 2 Context: EMAT Online Lesson Analysis for High School Teachers**

*a. What is the context and target population for this study, and how is STeLLA adapted for this audience?*

Building on the theoretical and empirical foundations of STeLLA, *Energy: A Multidisciplinary Approach for Teachers* (EMAT) is an online, 10-week graduate course that offers science content and pedagogical content experiences for in-service high school teachers. The overarching goal of the course is to support teachers in understanding how concepts about energy cut across science disciplines and to support them in teaching these concepts. The STeLLA approach of weaving science content and pedagogy experiences together through lesson analysis is emphasized. The EMAT course content is built around big ideas about energy that cut across disciplinary boundaries (physical science, biology, chemistry, physics, earth science, and environmental science). Throughout the course, teachers learn about energy concepts and engage in modified STeLLA PD experiences. The STeLLA PD was modified for the high school teacher audience, shortened to fit into the online course, and restructured for online delivery.

Unlike STeLLA, EMAT has no true face-to-face component; it is a facilitated online course that includes both asynchronous experiences and synchronous video analysis discussions. EMAT also targets a different audience. While STeLLA, ViSTA Plus, and RESPeCT all target elementary teachers, EMAT targets high school teachers. Because of the limited contact time for teachers in the EMAT course, we used a subset of the STeLLA teaching strategies—those that showed promise for impacting student learning (see yellow highlights in figure 8).

Each of the six EMAT course units (Coal, Nuclear Energy, Wind, Geothermal Energy, Biofuels, Solar Energy) has both a science content deepening strand as well as a STeLLA Lesson Analysis strand—both
of which were part of the STeLLA PD. In each unit, teachers participate in one synchronous discussion group where they analyze video and discuss pedagogical strategies to improve student learning. This abbreviated version of STeLLA PD included 12 hours of online, face-to-face PD in the synchronous discussions. Teachers were expected to and reported that they spent 10 to 14 hours per week to complete the 10-week course for a total of, on average 120 contact hours. About 48 of those hours were spent on lesson analysis work (40% of the course). In comparison, teachers participating in the STeLLA PD experienced 88.5 face-to-face contact hours across a one-year period (57 hours of lesson analysis and 31.5 hours of content deepening).

Learning to analyze science teaching through two lenses allows you to learn and use strategies for more effective science teaching

<table>
<thead>
<tr>
<th>Science Teaching</th>
<th>Strategies to Reveal, Support, and Challenge Student Thinking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1. Ask questions to elicit student ideas and predictions</td>
</tr>
<tr>
<td></td>
<td>2. Ask questions to probe student ideas and predictions</td>
</tr>
<tr>
<td></td>
<td>3. Ask questions to challenge student thinking</td>
</tr>
<tr>
<td></td>
<td>4. Engage students in interpreting and reasoning about data and observations</td>
</tr>
<tr>
<td></td>
<td>5. Engage students in using and applying new science ideas in a variety of ways and contexts</td>
</tr>
<tr>
<td></td>
<td>6. Engage students in making connections by synthesizing and summarizing key science ideas</td>
</tr>
<tr>
<td></td>
<td>7. Engage students in communicating in scientific ways</td>
</tr>
</tbody>
</table>

**Figure 8. Strategies Addressed in the EMAT Course**

b. **What are the goals for depth of impact, and how are we studying depth of impact?**

Coburn (2003) describes that for reforms to be “at scale,” they must affect deep consequential change in classroom practice. A goal of the EMAT project is to enhance teacher content knowledge, PCK, and practice related to teaching energy concepts in order to ultimately improve student learning. We are still analyzing the data from the second field test, but we have analyzed data from the first field test. To determine if EMAT teachers improved their science content knowledge, teachers were administered a pretest and a posttest to assess science content knowledge related to three big ideas in energy. These big ideas are as follows:

- No energy transformation is 100% efficient. Some energy always leaves the system as heat.
- Energy can be transferred or transformed but is never created or destroyed.
• Tracking fluxes of energy and matter into, out of, and within systems helps one understand the systems’ possibilities and limitations.

The test items were grouped according to these three big ideas and we analyzed the results by each major idea. Teachers (n = 25) in the first field test made significant gains in their content knowledge from pretest to posttest. Figure 9 graphically illustrates the changes from pretest to posttest by each science idea. The data associated with the table indicate the statistical significance of each change, the effect sizes for the change, and in brackets, the confidence interval for each effect size.

![Figure 9](image)

Figure 9. Teachers’ posttest scores were significantly better than their pretest scores for each key idea.

EMAT teachers also took a pretest and posttest to determine how well they could apply what they had learned to analyzing videos of science teaching. They analyzed video from four classroom teachers. BSCS researchers analyzed their written analyses to determine if participants became more analytical about issues related to the Student Thinking Lens (STL) and the Science Content Storyline Lens (SCS). The pre/post differences were significant at p < .001 (both on the total test score and on the SCS and STL individual pieces). The pre/post effect sizes are big and comparable to what Roth and colleagues (2011) found in the first study of the STeLLA face-to-face PD program (table 3).

<table>
<thead>
<tr>
<th>Subscale</th>
<th>STeLLA (2011) pre/post effect size</th>
<th>EMAT pre/post effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>STL items</td>
<td>1.31</td>
<td>1.16</td>
</tr>
<tr>
<td>SCS items</td>
<td>1.07</td>
<td>1.58</td>
</tr>
</tbody>
</table>
EMAT teachers video recorded a lesson both before taking the course and after they completed the course. We are in the process of analyzing this video for evidence of their use of the STeLLA strategies to make student thinking more visible and to create coherent content storylines in their lessons. We will ultimately use the video scores as covariates in HLM models to predict student achievement. We anticipate that teachers’ post videos will show improved use of the strategies. The results of the first field test showed that teacher effects did not “trickle down” to students. There was no significant difference between groups on student achievement. Because of these results, we modified the program to try to improve student learning (additional units with lesson analysis, more emphasis on the big ideas related to energy, more attention to challenging teachers to apply the big ideas to their everyday teaching, and so forth). We have not analyzed student data from the second field test to see if these modifications were successful.

Through the EMAT study, we will be able to determine if the STeLLA approach to PD can be effective in an online format with high school teachers. We have already shown that teachers can make impressive gains in content knowledge and in their ability to analyze classroom video for the use of the STeLLA strategies. What remains to be seen is the extent to which the online course can also support transformations in teacher practice. Furthermore, the second field test results will provide additional information about the extent to which the online materials for teachers can lead to improved student learning. A key difference between EMAT and STeLLA is that EMAT lacks curriculum materials for students to use. It may be that the presence of curriculum materials to support student learning and to serve as a model for teachers to develop their own STeLLA lessons may be a critical component to the STeLLA PD program.

c. How will this study help us learn about the potential for spread of the STeLLA PD approach, including the spread of ideas (such as the STeLLA conceptual framework and its teaching strategies)?

The EMAT study is interested in answering the question: Will the STeLLA PD approach be effective with high school teachers delivered as an online course? The answer to this question will help us understand the more traditional idea of spread—spread to a wider audience. Taking the STeLLA teaching strategies to a new audience will have the potential for the reform to reach greater numbers of classrooms and schools.

However Coburn (2003) describes spread as going beyond greater numbers of classrooms and schools to include spreading reform-related norms and pedagogical principles within a classroom, school, and district. Will high school teachers continue beyond the study to incorporate the STeLLA strategies into their own planning and teaching? And will they share their experiences and STeLLA knowledge with colleagues and with their district? To address these questions, we are conducting a survey that asks teachers about these aspects of spread defined by Coburn. We have anecdotal evidence that some teachers are sharing with their colleagues and using the strategies in other classes and in other units of study. One teacher structured her course syllabus to be organized around the big ideas of energy. She also provided PD for her district based on the STeLLA teaching strategies. We will use the results of the survey to determine if EMAT teachers are sharing and incorporating reform-based norms and pedagogical principles beyond the lesson they video recorded. This information will help us determine the potential for spread of these reform-related norms and pedagogical principles within a classroom, school, or district.

d. How does this study explore/address issues of sustainability of the reform effort?

Coburn (2003) places the idea of sustainability as a vital component of true scale. When considering scale and sustainability, the two should not be treated independently but rather should show how scale depends
on sustainability. For a program to be truly scalable it must also be sustainable. Coburn describes sustainable reform efforts to be ones that provide strategies and tools that teachers, schools, and districts will need to sustain the reform after the initial influx of resources dissipates.

The EMAT study addresses issues of sustainability minimally. In retrospect, we could have paid more attention to sustainability issues from the beginning and in the research design. The EMAT curriculum was tested through Montana State University’s National Teachers Enhancement Network (NTEN). We are currently exploring whether it will be sustainable for NTEN to continue to offer EMAT for graduate credit after the grant period ends. Can they continue to support the synchronous online lesson analysis component? We will also learn from our teacher surveys if teachers are sustaining the STeLLA approach to PD and/or the STeLLA approach to teaching without the resources of BSCS or the course support. This will only tell us if the teachers continue to use the strategies or use the resources for PD in the year following the course. To understand if the strategies are truly sustainable beyond the life of the project, measures must be built in to the initial project to provide teachers the tools to promote sustainability.

e. Is there a plan for shifting ownership of the reform to the schools, districts, or university partners?

Shift in reform ownership is another facet of scale described by Coburn (2003). The STeLLA strategies for teaching and the approach to PD initially were “owned” by BSCS. EMAT teachers who are incorporating the strategies into their everyday teaching or sharing the strategies and approaches with colleagues have begun to shift the ownership from the BSCS researchers to themselves. However, EMAT participation is currently a decision made by individual teachers; therefore, “ownership” of ideas is limited to the individual teachers. The shift in ownership would be strengthened if districts adopted the STeLLA approach to PD or if EMAT teachers shared the approaches and strategies with their colleagues in district or school PD experiences. For the reform effort to be truly scalable, the ownership must shift from the researchers to the teachers and ideally to the schools and districts.

The EMAT course provides tools and strategies to help individual teachers take ownership of the STeLLA teaching strategies in their teaching of energy concepts. For example, in the synchronous lesson analysis discussions, teachers are analyzing their own and each other’s teaching. In planning and teaching these lessons, teachers are prompted to use knowledge about the energy concepts and the STeLLA strategies. During synchronous lesson analysis discussions, they are encouraged to talk with each other about ideas for improving their lessons. In addition, they are prompted to share ideas of incorporating the teaching strategies more broadly in their classes. This course scaffolding supports teachers in transferring the ownership of the strategies and approaches from BSCS researchers to the teacher.

Our dissemination plans are to offer the course for free from the BSCS website. The course could be used by individual teachers in a nonfacilitated format, or district PD leaders will have access to a facilitation guide to support them in working with teachers either in online synchronous discussions or in face-to-face discussions. This plan has the potential to shift ownership away from the BSCS researcher. However, the ownership of the course will only shift to schools, districts, or even the individual teacher if these entities use the course to meet their own needs and their own goals.

f. To what extent do students, teachers, school districts, and researchers who are involved in the project share common goals and support activities related to the success of the reform effort? What are we learning about competing goals and activities across these groups that threaten the success of the reform effort?

Drawing from the work of Lynch (2012), when all relevant activity systems are engaged and their goals and support activities are harmonious, conditions exist for an intervention to go to scale. We have identified four main activity systems needed to take the EMAT reform effort to scale: researcher, teacher,
The EMAT reform effort actively engages the teacher and researcher activity systems but largely ignores the student and school district activity systems. The goals of the teacher and researcher activity systems are mostly congruent and support the reform, but competing goals did arise. Teachers applied for the EMAT project based on a match between the course goals and the teachers’ desire to improve their science content knowledge, pedagogical content knowledge, and practice. However, it became apparent during the course that some teachers did not see the value in studying energy in varied contexts—those beyond the courses they were currently teaching. Unlike the elementary teachers in STeLLA, high school teachers are typically teaching only one discipline of science and are looking for ideas, strategies, and approaches to enhance their teaching in that discipline. The online EMAT course includes science content related to many science disciplines, including life science (through the biofuels unit), physical science (through the coal, nuclear, and solar energy units), and Earth science (through the wind and geothermal heat exchange units). We tried to create coherence by emphasizing the three big energy ideas (described above) in each unit. In spite of the fact that all six units emphasized the same three crosscutting ideas, and that all three crosscutting ideas are important in all science disciplines, some teachers expressed concern that they would never use the content from the course that was outside their area of expertise. This implies a conflict of the course goal of improving the teachers’ science content knowledge related to crosscutting ideas and the teachers’ goal to improve content knowledge within the confines of their own discipline. From this conflict, we can learn to be more explicit in the course goals as we move forward and disseminate the EMAT course to more teachers. We can also consider alternative models for helping high school teachers understand crosscutting ideas when teachers are unaccustomed to thinking in a cross-disciplinary way.

The EMAT study did not significantly engage the students or the district. Students were involved only to the extent that they were given pretests and posttests to determine their gains in science content knowledge related to the three big ideas around energy. The only engagement that the researchers had with school districts was to get permission for video recording and student testing. This lack of foresight and attention to meaningful engagement of the school district and student activity systems compromised the chance that the EMAT reform effort would be scalable. In future projects, engaging the district may help the reform to spread to other teachers in the district, promoting the scalability of the effort. For example, researchers might seek to work with district science coordinators alongside the interested teachers to promote communication among all of the partners. In addition, BSCS could create and disseminate briefs describing the project and the results with district personnel.

Study 3 Context: RESPeCT Teacher Leaders and District-wide Implementation

a. What is the context and target population for this study, and how is STeLLA adapted for this audience?

A major challenge that is highly relevant for school districts and the science education research community is how to implement rigorously tested science education professional development and practices in a manner that is practical and sustainable. In a National Science Foundation-funded partnership project, California State Polytechnic University—Pomona (CPP), BSCS, and Pomona Unified School District (PUSD) are studying whether the STeLLA PD approach can be scaled to eventually reach all K–6 teachers in a high-needs, Title I school district. The key to this scaling-up effort is the development of a cadre of grade-level teacher leaders who can lead the PD with their peers.

1 PUSD is a district in eastern Los Angeles County that serves an ethnically diverse population of over 33,000 students enrolled in 41 schools. The majority of students in grades 5 and 8 scored between “Basic” and “Far Below Basic” on the 2010–11 Science California Standards Test. In PUSD, 88% of students are Hispanic, 7% are African American, and 39% are English learners.
The project, *Reinvigorating Elementary Science through a Partnership of California Teachers* (RESPeCT), is modifying the original SteLLA PD program to better support scale-up and sustainability within the district. The target audience is now all elementary teachers, grades K–6. The original SteLLA work focused on grades 4 and 5. This impacts the eventual reach of the program. The program will be adapted to include videocases, lesson plans, PD Leader Guides, and other materials for grades K, 1, 2, 3, and 6 in addition to grades 4–5. As well, the intent is to reach all K–6 teachers in the district, not just volunteer, early adopters.

To support the development of teacher leaders, the project is also designing, implementing, and studying a RESPeCT leadership program. At each grade level, six teacher leaders will be prepared to learn to lead the PD program with their peers. In their first year, these teachers will be participants in the year-long RESPeCT PD program which will replicate the SteLLA PD program (Roth et al., 2011) with modifications designed to increase its sustainability (see section d below). The program will be co-led by experienced BSCS SteLLA PD leaders and CPP science and math faculty. During year 2, the teachers will participate in a year-long RESPeCT Teacher Leadership program. During the leadership summer institute, they will study the RESPeCT PD Leader Guide, learn about effective PD practices, and practice leading lesson analysis and content deepening PD sessions. During the school year, they will again teach the RESPeCT lesson plans, but this time they will learn how to select the video clips and how to lead the lesson analysis study group meetings. In year 3, they will work in pairs to lead the PD program for a group of their peers.

b. *What are the goals for depth of impact, and how are we studying depth of impact?*

The goals for teacher and student learning in RESPeCT parallel those goals in SteLLA II: improved teacher science content knowledge, teacher pedagogical content knowledge, teaching practice, and student science content knowledge. The project will use similar measures as in SteLLA II to assess teachers’ content knowledge, teachers’ pedagogical knowledge, and students’ science content knowledge. Unfortunately, cost issues prohibit the analysis of videotapes of teachers’ practice that was done in the SteLLA II study. The study uses carefully selected pairs of matched schools to allow for comparison of student learning in treatment and control schools.

During the life of the grant, 168 teachers and over 5,040 students will participate in the RESPeCT PD (30% of PUSD elementary teachers). By the end of the RESPeCT grant, the district will have the self-sustaining capacity to continue to implement a program for science PD that has a direct, positive impact on student learning. The direct involvement of principals and the Assistant Superintendent throughout the RESPeCT project along with PUSD’s commitment to use Title II funds to keep the program going after the life of the grant ensures the program will ultimately reach all 551 elementary PUSD teachers.

c. *How will this study help us learn about the potential for spread of the SteLLA PD approach, including the spread of ideas (such as the SteLLA conceptual framework and its teaching strategies)?*

We know that the SteLLA PD is effective for teachers and students when implemented with volunteer teachers from a variety of school districts and when led by BSCS PD leaders. We do not know whether such high-quality, research-based PD can have the same effects when led by school district teacher leaders and when implemented district-wide in a high-needs school district. This study is not only testing a modified SteLLA PD program in a district-wide context, it is also producing and studying a Teacher Leadership program. This program includes a rich array of materials to support teacher leaders, including PD Leader Guides and Powerpoints, videocases, and lesson plans. These supports will add to the potential spread of this approach to other districts beyond the life of the grant.
We also anticipate that the project’s investment in the CPP science faculty will lead to the spread of this PD approach to other districts. Under the mentorship of experienced BSCS PD leaders, the CPP faculty are developing expertise in leading analysis-of-practice professional development. The goal is that they will eventually have the skills and the commitment that will enable them to take the lead in preparing RESPeCT teacher leaders in other school districts. That is, they will have the capacity to lead teachers through a two-year development process: 1) year 1 participation in the RESPeCT PD program and 2) year 2 participation in the RESPeCT Teacher Leadership program.

It is going to be the student learning results that will be the most powerful determinant of the potential spread of this program. Very few studies of science PD exist that examine the impact on student learning, and even fewer explore the impact on student learning at a district level and longitudinally. If this study can demonstrate impact on student learning when the program is led by teacher leaders, this will make the program attractive to other districts and will greatly increase the potential for spread. Without these data, it would be challenging to attract district interest in implementing an approach that is not quick, easy, or cheap.

d. How does this study explore/address issues of sustainability of the reform effort?

The STeLLA PD program has been modified in a number of ways that we hope will increase its ability to be sustained over time in high-needs school districts:

- The target science content and science practices will closely match the NGSS. Match to NGSS will better support district buy-in and strengthen students’ opportunities to demonstrate their learning on state assessments.
- The STeLLA strategies are modified to more closely match the NGSS way of defining science practices.\(^2\)
- The STeLLA lesson plans that teachers analyze and teach will integrate and highlight Common Core English Language Arts (ELA) and Math standards. This will support better district and teacher buy-in and thus support wider spread of the program and sustainability beyond the life of the project.
- The lesson plans used in the program are being revised to more strongly and more explicitly address the needs of English language learners. Thirty-nine percent of Pomona students are English learners.
- The last semester of the year-long PD program will be restructured to provide teachers with lesson plans that highlight the STeLLA strategies instead of engaging teachers in planning their own lessons. This revision will enhance the likelihood that teacher leaders can implement and sustain the program successfully.
- The program will be modified to more closely integrate the content deepening activities with the STeLLA lesson plans and lesson analysis work. This is being done for two reasons. First, results from the STeLLA II study show that teachers’ content learning was enhanced by their participation in lesson analysis work. Secondly, STeLLA I teachers had access to science content specialists; in RESPeCT, teacher leaders will be the main support for teachers’ science content learning. By integrating the content deepening even more closely with the lesson plans and lesson analysis work, we are creating a program that can be led by teacher leaders.
- Professional Development Leader Guides and Powerpoints are being developed to support teacher leaders in enacting the RESPeCT PD with their peers.

---

\(^2\) Previous STeLLA Strategy: Engage students in interpreting and reasoning about data and observations. Revised STeLLA Strategies: Engage students in analyzing and interpreting data and observations. Engage students in constructing explanations and arguments.
e. **Is there a plan for shifting ownership of the reform to the schools, districts, or university partners?**

The RESPeCT program explores two key challenges of sustainability: 1) capacity building of content faculty who can support elementary teachers and 2) capacity of a high-needs district to continue to use a video-based analysis-of-practice PD model. This reform effort is designed with the intention that over time, ownership of the STeLLA PD approach will shift from BSCS to CPP and to PUSD teachers, schools, and central administrators.

**CPP ownership.** In the first two years of the project, experienced STeLLA PD leaders from BSCS mentor CPP science faculty in becoming effective leaders of a lesson analysis approach to PD. Through workshops and through co-planning and co-leading the lesson analysis PD, BSCS PD leaders support CPP science faculty in taking on this role. The goal is for CPP science faculty to develop the capacity to contribute to long-term impact through work with other districts after the life of the grant. In addition, a goal is that CPP science faculty will use the STeLLA framework and strategies in their own planning and teaching of university science courses for preservice elementary teachers. To reach these two goals, it is essential for CPP science faculty to take ownership of the analysis-of-practice professional development role and of the STeLLA conceptual framework and strategies.

What structures are in place to encourage science faculty sustainability and ownership of the effort?

- This effort aligns with the CSU Chancellor’s commitment to doubling the number and quality of science and mathematics teachers in California.
- CPP faculty are given course release time and summer pay to participate in project activities.
- CPP faculty are encouraged to contribute to publications about this work, thus enhancing their professional curriculum vitae.
- CPP faculty are members of design teams that construct lesson plans for use in the PD program. They play a central role in contributing their content expertise in this work.
- CPP faculty take the lead in designing and implementing content deepening activities for teachers in the summer institutes.
- As they co-lead lesson analysis sessions with BSCS PD leaders, CPP faculty are supported in taking on the PD leadership role; as they are ready, they gradually take on increasing leadership of these sessions.

**PUSD ownership.** Success of this reform effort depends on PUSD administrators and teacher leaders taking ownership of and sustaining the RESPeCT PD program/approach. Key to this ownership will be the student learning results gathered during years 3, 4, and 5 of the research study. The administration has made it clear that they can commit to supporting the professional development program beyond the life of the grant, but their level of commitment will obviously be highly influenced by the student learning results.

In addition to the student learning results, what other structures are in place to encourage school district ownership of the effort?

- Common Core Math and ELA standards are integrated into the RESPeCT PD lesson plans. If teachers and administrators see the science work and the STeLLA strategies as useful in reaching curriculum goals in math and ELA, this will enhance the opportunity for buy in.
- Through participation in this PD, teachers will become eligible for the new Foundational Level General Science credential, increasing their job security and employment options.
- Principals are involved in the substance of the program. They attend workshops about the program, and they observe some summer institute sessions. They know what teachers are learning and are supportive of efforts to use that learning in their teaching.
• Support from the Superintendent’s office communicates the value of this work to teachers. The Assistant Superintendent is a Co-PI on the grant. Both the Superintendent and the Assistant Superintendent attend portions of the summer institute.
• ELA and elementary science teacher specialists who normally provide PD for PUSD teachers are involved in designing the lesson plans that are used in the program, and they are participating as full participants in experiencing the year-long RESPeCT program. This readies them to support the teacher leaders.
• PUSD is poised to be successful with this kind of PD because of past experience with teacher learning teams that focused on improving student achievement in ELA and math. A survey in 2010 of 89 teachers suggested that 97% observed positive results in student performance due to the program. We predict that there will be strong buy-in from PUSD teachers for the RESPeCT PD approach.

f. To what extent do students, teachers, school districts, and researchers who are involved in the project share common goals and support activities related to the success of the reform effort? What are we learning about competing goals and activities across these groups that threaten the success of the reform effort?

At this point, midway through the second year of the project, all partners—BSCS, CPP faculty, PUSD administrators and teacher leaders—are supportive of the goals of this partnership. All partners agree that an unacceptable number of K–6 students in California are failing to develop a rudimentary understanding of science, much less the more advanced understandings described in the Next Generation Science Standards and that there are unacceptable achievement gaps that must be addressed. In addition, there is support that the lesson analysis approach to teacher professional development is a good way to tackle these problems in Pomona elementary schools. The partners are working hard to build trust and strong collaborative relationships. The biggest challenge we have identified so far is that each activity system—teachers, administrators, university faculty, researchers, and BSCS curriculum writers—has competing activities making substantial demands on their time so that it is difficult to find common meeting times and to complete tasks on time. As the project progresses and teacher leaders begin leading the program with their peers, we anticipate we will have much more to say about the various activity systems in play and the degree to which they are in sync or in competition.

Summary. These three projects—VISTA Plus, EMAT, and RESPeCT—are exploring different ways in which the STeLLA videocase-based, analysis-of-practice approach can be adapted to reach additional audiences and to support the scalability and sustainability of the program beyond the life of research grants. Using quasi-experimental research designs, each of the projects looks at impact on both teacher knowledge and student knowledge. In addition, both VISTA Plus and EMAT examine videotapes to assess changes in teaching practice. While additional findings will emerge from these projects, it is already clear from the results of the STeLLA I and STeLLA II studies that the professional development approach has significant impact on teacher and student science learning. What are the features of the PD that contribute to this success? In the next section, we explore features of the STeLLA PD that expand the current consensus view of effective PD.
Part 4: Findings from the STeLLA Line of Research Expand the Consensus Model of Effective PD

Professional development is an important mechanism for increasing teacher knowledge and changing teaching practice. However, the field is still working to identify the underlying mechanisms that make certain professional development programs effective (Wilson, 2013, Reiser, 2013). As reported in Parts 1 and 2, the STeLLA PD program demonstrated significant impacts on teacher and student science learning in two rigorous studies. What were the key features or mechanisms of the PD that enabled teachers to support such student growth? In Part 1 of this paper, we described the key features, substance, and form of the STeLLA PD program. In this section, we highlight how the STeLLA program aligns with features of the consensus model of effective PD (figure 1). We then consider STeLLA program features that push beyond the consensus model and nominate these features as a starting point for building a new research-based model of PD that is backed by evidence of impact on student learning.

**Design/procedures.** The research team analyzed the features that characterize the STeLLA PD program. Some of these features focus on the program form (e.g., duration, collaboration in study groups), while others characterize the substance of the PD program—the ideas, strategies, and practices that teachers are learning from the program. From this list of program features, we identified a subset that were not captured in the reviews of research on teacher PD as being part of what researchers now describe as the consensus model of effective PD. We then examined survey and interview feedback from STeLLA teacher participants and the STeLLA professional development leaders regarding the program features that they pointed to as particularly influential. Through this process we identified six key features of the STeLLA PD model that were not identified in the consensus model as described by Desimone, Porter, Garet, Yoon, and Birman (2002), Desimone (2009), Garet et al. (2001), and Wilson (2013).

**Findings, Part 1: STeLLA and the Consensus Model**

The STeLLA program was designed by the principles listed in figure 3. These principles overlap with much of what has emerged consistently as features on a consensus model of effective professional development. Ways in which the STeLLA PD program matches and does not match the consensus model features are described next.

**Focus on Science Content.** During their summer institute, STeLLA teachers attend a science content deepening session in the morning and a lesson analysis session in the afternoon as part of the summer Lesson Analysis PD program. Both sessions focus on learning specific science content in their grade-level curriculum. In the science content sessions with university professors, teachers engage in inquiry-based science experiences as adult learners. During the lesson analysis portions of the day, teachers critically examine classroom videos of teachers teaching science concepts and ideas that were the focus of the morning content deepening session with the purpose of analyzing student thinking that was revealed, supported, and/or challenged through the use of the STeLLA teaching strategies. Throughout this process, teachers are continually required to reflect upon and share their own ideas concerning the science content. This leads to opportunities to clarify science content questions, confusions, and misunderstandings.

**Active Learning.** Throughout the STeLLA program, teachers apply their science knowledge and understanding of the STeLLA framework and strategies to actual teaching contexts. During the summer institute, teachers are immersed in inquiry-based experiences that required them to make sense of natural phenomena through the use of explanatory models. Likewise, during the lesson analysis component of the summer program, teachers analyze videocases that showcase teachers teaching science to students and that provide opportunities to consider how the use of the STeLLA strategies contributed, or could have contributed, to student learning of science content. During the
academic year, teachers apply their knowledge and understanding of the science content and STeLLA strategies to their own classroom science planning and teaching. In the fall, they analyze their teaching of the STeLLA lessons and examine work from their own students. In the winter/spring, they are actively and collaboratively involved in planning lessons in a new content area, using the STeLLA framework and strategies as a guide. Thus, throughout the program, teachers are actively engaged learners.

**Program Coherence.** Although we believe the STeLLA program to be coherent, we describe that coherence in a different way than is described in the literature. Coherence as described in the PD consensus model refers to “the extent to which teacher learning is consistent with teachers’ knowledge and beliefs” as well as the degree to which the content of the professional development is consistent with school, district, and state reforms and policies (Desimone, 2009, p. 184). STeLLA is aligned with national and state reform documents and standards, and it builds on teachers’ prior knowledge and beliefs. However, we believe that the coherence of the STeLLA program is far different than is described in literature supporting the consensus model where coherence is based on alignment with teachers’ prior knowledge and beliefs and school district initiatives. For example, the STeLLA program has components that are designed specifically to challenge teachers’ prior knowledge and beliefs. We will discuss the coherence of the STeLLA program as different from the consensus model in our discussion in “Findings, Part 2,” where the ways in which the STeLLA program pushes beyond the consensus model are described.

**Sufficient Duration.** Research results concerning effect sizes described in part 2 of this paper provide evidence supporting the assertion that the STeLLA PD program was sufficient in both duration and intensity. The STeLLA program was a yearlong program starting in early summer with a two-week summer institute followed by eight 4-hour study group sessions during the academic year. The summer institute included 27 hours of lesson analysis and 31.5 hours of content deepening for a total of 58.5 hours of professional development. The academic-year study group sessions included 28 hours of lesson analysis with a culminating 2-hour session at the end of the academic year for a total of 30 additional hours of professional development. The total number of hours teachers participated in the professional development program was 88.5 hours. The program thus had not only substantial total amount of contact hours, but teachers’ learning was distributed across an entire year and included additional time they spent working on their planning and teaching outside of PD meetings.

**Collective Participation of Teachers.** This is defined in the literature as involving all the teachers from the same school, grade, or department. This sets up more potential interactions and discourse related to the PD as teachers in the same setting experience the same professional development. Because they encounter each other on a daily basis, teachers have more opportunities for talk and action related to PD goals but taking place outside the formal professional development meeting times. The STeLLA program gathers teachers together in study groups of the same grade level, but in both the STeLLA I and STeLLA II study, participating teachers came together from a variety of schools and districts. We rarely had more than two teachers from the same school. For example, in STeLLA II, there were 144 teachers involved, and they came from 77 different schools spread across the Colorado Front Range.

**Findings, Part 2: Beyond the Consensus Model**

Does the current consensus model provide a good framework for PD design and development? Does it capture key features of PD design that make a difference in student learning? Questions about its effectiveness remain. According to Hill, Beisiegel, and Jacob (2013), disappointing results from recent rigorous studies of programs containing some or all of these features provide sufficient reason to question whether this consensus model goes far enough to define the elements of effective PD for science teachers.
Testing the Consensus Model of Effective PD

(Arens et al., 2012; Bos et al., 2012; Garet et al., 2008; Garet et al., 2011; Santagata et al., 2011). In this section of our findings, we describe the features of the STeLLA PD that are not highlighted in the consensus model but that were identified by STeLLA PD leaders and participants as influential in obtaining the teacher and student learning gains described in this paper.

There are several features of the STeLLA program that participants and PD leaders identified that go beyond the consensus model and that might help to explain the program’s positive impact on teacher and student learning. Some of these features have been previously identified in the literature, especially in conceptual pieces that describe what teachers’ professional learning should look like (e.g., Ball & Cohen, 1999; Thompson & Zeuli, 1999). However, these features were not supported by enough empirical evidence to be included in the currently-accepted consensus model. Other features appear to be unique to the STeLLA program and may introduce new ways of thinking about the features of effective professional development.

Features of the STeLLA program we identified that went beyond the consensus model of effective professional development include the following:

a) The program substance is organized around a conceptual framework.
b) Science content learning is intertwined with analysis of practice.
c) Analytical tools and videocases support collaborative, deep analysis of science teaching, student learning, and science content.
d) The professional development program is guided by a theory of teacher learning.
e) There is an internal coherence of program form and substance.
f) Learning is directed and scaffolded by knowledgeable PD leaders.

a. Program substance is organized around a conceptual framework. In the consensus model of effective PD, the main comment about program substance is that it should focus on specific subject matter content. In contrast, the substance of the STeLLA PD went beyond the science content and was tightly organized around a conceptual framework that specified what teachers would learn about pedagogy and pedagogical content knowledge. The STeLLA framework consists of two lenses, each shaped by the research literature on effective science teaching and learning: (1) The Science Content Storyline Lens describes nine discrete planning and teaching strategies to focus science lessons on the science ideas (rather than simply fun activities or procedures) and to help students construct connected understandings as the “story” unfolds from data and other experiences; and (2) the Student Thinking Lens identifies eight strategies intended to help teachers reveal, support, and challenge student thinking. Recent research on changing teacher practice recognizes the importance of focusing on such “high-leverage” practices (Ball, Sleep, Boerst, & Bass, 2009; Smith & Stein, 2011; Windschitl et al., 2012). The STeLLA framework intentionally shifts the focus of the STeLLA PD from simply the science content to be taught and ways of delivering that content to focus on how students are making sense of the science content throughout a course of instruction. The lenses and strategies focus teachers’ attention on students’ incoming ideas about science content and effective pathways to move students from their current explanations of the world to more scientific ways of thinking, a feature of PD often mentioned in the literature but not included in the consensus model (Kennedy, 1998; Ball & Cohen, 1999; Borko, 2004; Hawley & Valli, 2006; Vescio, Ross, & Adams, 2008; Guskey & Yoon, 2009; Wayne, Yoon, Zho, Cronen, & Garret, 2008).

b. Science content learning is intertwined with analysis of practice. The consensus model of PD suggests that effective professional development should focus on specific content, and in most cases,
this is described as the science content being taught in the classroom. Additionally, much of the research literature on PD indicates the need for critical analysis of teaching and reflection with peers using common artifacts of practice (Wilson & Berne, 1999; Ball & Cohen, 1999; Borko, 2004, Hawley & Valli, 2006; Vescio, Ross, & Adams, 2008). One of the unique aspects of the STeLLA program is the way science content learning is married to the critical analysis of practice through the artifacts of common classroom lessons and videos. Ball and Cohen (1999) describe the need for these two elements of the professional development experience to be closely intertwined. Our experience in the STeLLA program would support the concept that teachers need close connections between the content focus of PD and the critical analysis of teaching of that content. In a measure of science content knowledge, teachers in the STeLLA I Lesson Analysis program outperformed teachers in the Content Deepening program (see results in part 2 of this paper) despite the fact that teachers in the Content Deepening program experienced twice as many hours of content-specific professional development. This result matches both Lesson Analysis teachers’ and PD leaders’ perceptions that the STeLLA lesson analysis work provided a meaningful context for deepening teacher understanding of the science content. It was through discussions of specific lesson content in videocases that teachers deepened their own content understanding in ways that transferred to their classroom teaching. These discussions often focused on common student ideas heard during their own or others’ classroom teaching situations and ways these ideas continued to influence student thinking despite experience and instruction. Teachers often realized that they held similar misconceptions that had not been addressed in their content work with university science instructors. Throughout the lesson analysis sessions, teachers not only identified the science ideas expressed by students during classroom instruction but also actively discussed accuracy of student ideas, shared their own knowledge and understanding of the science ideas and concepts, and debated what strategy or strategies would be useful to support and/or challenge student thinking.

c. Analytical tools and videocases support collaborative, deep analysis of science teaching, student learning, and science content. The STeLLA PD provided a variety of tools and materials that supported teachers in deep analysis of science teaching, student learning, and science content. The videocases showcasing experienced teachers provided participating teachers with images of the STeLLA strategies being used at their specific grade level and with the specific science content they are teaching. The videos provided opportunities for teachers to identify effective use of STeLLA strategies as well as missed opportunities. A STeLLA Lesson Analysis Protocol (figure 5) supported teachers in analyzing video clips from their own and their peers’ classrooms during the school year. A Features Analysis Chart (figure 6) enabled teachers to assess particular strengths and weaknesses in their students’ written work. The use of these and other analytical tools had a number of positive impacts on teacher learning. First, this structure helped teachers know what to expect and what was expected of them, thus contributing to their sense of safety in the learning community. Additionally, these tools provided structures that guided discussion and analysis in productive and focused pathways. The intentional selection of video clips and analysis questions that were focused on the STeLLA student thinking and science content storyline strategies also helped keep discussion and analysis from straying into unproductive or ancillary topics and helped teachers focus on the heart of the STeLLA analysis—student learning. Ball and Cohen (1999) mention the importance of analytical tools to support teacher learning from practice, but this is a feature of effective PD rarely highlighted in the literature.

In addition to the videocases and analytical tools, teachers are provided with detailed lesson plans that highlight how the STeLLA strategies can be integrated into science teaching practice. In their analysis of video clips from this teaching, teachers focus on their efforts to implement particular STeLLA strategies and also identify missed opportunities to use them most effectively. PD programs often assume that translating new ideas into practice is a straightforward process; the STeLLA model
recognizes the challenge of changing one’s teaching practice and provides opportunities and experiences that support that change.

d. **A professional development program guided by a theory of teacher learning.** Having a foundational theory of teacher learning and change is cited by both Ball and Cohen (1999) and Yoon et al. (2007) as an important element of effective PD. The STeLLA program is grounded in a situated cognition theory of teacher learning and a cognitive apprenticeship instructional model (Lave & Wenger, 1999; Collins, 2006). There are three program elements that embody this theory of teacher learning: 1) a clear learning progression beginning with highly scaffolded learning experiences and moving toward greater levels of independence, 2) experiences that create in teachers a “need to know more” by creating a level of cognitive dissonance and opportunities to resolve that dissonance with alternating cycles of classroom implementation and collegial analysis and reflection, and 3) the use of experts who plan for and guide teacher learning.

The program has a clear learning progression, starting with a highly scaffolded summer institute where participants are introduced to the strategies and learn how to analyze teaching and learning through videocases from other teachers’ classrooms. In study group work across the school year, this scaffolding is gradually reduced. In the fall, participants are scaffolded in using STeLLA strategies by implementing prepared lessons in their own classrooms and by analyzing their own and peers’ practice. In the spring, teachers collaboratively develop lessons incorporating the STeLLA strategies with fewer scaffolds to guide enactment of strategies in instruction.

The learning sequence is designed to create a high level of cognitive dissonance for teachers as they wrestle with the science content and as they watch and analyze classroom teaching and compare these instances with their own teaching practices and beliefs about teaching and learning. Throughout the analysis of videos, teachers are able to negotiate new understandings, consider alternatives, and contribute to one another’s thinking. By using prepared lessons and analyzing video from this enactment, teachers are provided a means through which they can develop a new repertoire of skills and abilities to fit with their new understandings, and they take the time to consider decisions made during that enactment that match with, or do not match with, the STeLLA framework. As teachers collaboratively design their own lessons incorporating STeLLA design principles (from the Science Content Storyline Lens) and structural elements (from the Student Thinking Lens) in the second semester of the program, they have the opportunity to strengthen their understanding of the strategies and transfer it to their own planning and teaching practice. The inclusion of these dissonance-causing elements along with intentional ways of resolving dissonance and developing a new repertoire of practice are described by Thompson and Zeuli (1999) as essential elements of transformative teacher learning that support this component of the theory of teacher learning that guides the STeLLA program.

In the STeLLA PD program, several “experts” guide the learning of teachers as they learn about and then enact new understandings about teaching and learning. Several researchers have noted the importance of external experts to push teacher learning beyond accepted cultural norms (Ball & Cohen, 1999; Borko, 2004; Guskey & Yoon, 2009). In STeLLA, these experts include the PD leaders, university-level science content experts, as well as the teachers showcased in video clips who provide examples of the use of STeLLA strategies in grade-specific classroom situations. Each plays a role in the cognitive apprenticeship learning theory.

e. **Coherence of the PD program in form and substance.** Whereas the consensus model of PD describes the importance of coherence as being aligned with policies external to the program such as teachers’ prior knowledge and beliefs, reform documents, and school policies and practices, we believe a crucial element of the STeLLA PD program is the internal coherence of the program in both
form and substance. In terms of substance, the program is tightly focused throughout on the ideas and strategies embedded in the STeLLA conceptual framework and a small set of core science ideas in two topic areas at each grade level. The program is intentional in how it sequences learning experiences related to this content, with multiple opportunities for teachers to revisit and deepen understandings of key ideas. For example, teachers initially learn about the STeLLA strategies through reading and discussing program documents and analyzing video cases of other teachers enacting the lessons. Later they learn from teaching and analyzing their own enactment of lessons, and finally they engage developing lessons plans embodying the STeLLA lenses and strategies.

Internal coherence is also visible in the organization of individual PD sessions, each of which is carefully designed using the Science Content Storyline Lens—starting each meeting with focus questions, selecting activities matched to the learning goal, and building deeper understanding by explicitly linking science ideas and ideas about the strategies to each activity. PD leaders use Student Thinking Lens strategies during the sessions to reveal and challenge teachers’ thinking about science content as well as their beliefs about teaching and learning. Teachers analyze data from their classrooms and analyze video using an explanatory framework including claims, evidence, reasoning, and alternatives. They synthesize and reflect on new understandings. The teacher learning in the program is designed to allow teachers to experience the power of the STeLLA lenses and strategies not only for science content learning but also for their learning about science teaching.

f. **Learning is scaffolded by knowledgeable PD leaders.** While the role and the background of the PD leaders are rarely mentioned in discussions of effective PD (see Ball and Cohen [1999] and Borko [2004] for two exceptions), the STeLLA study analyzed the knowledge and skills of the PD leaders and found that part of the success of the STeLLA program lies in the role of knowledgeable PD leaders who create a safe environment for teachers to challenge their own understanding of teaching, learning, and science content. While leading sessions, STeLLA PD leaders draw from knowledge about the specific science content; knowledge about the STeLLA framework, strategies, videocases, and analytical tools (e.g., the PD curriculum); and knowledge about ways to support adult learners and to develop an environment where teachers are both supported and challenged to grow and change. We call this knowledge PCK for PD leaders (a concept that is more fully explored in Landes and Roth, 2013).

Because of the strong student learning results, the STeLLA program enables the science education community to move beyond a “consensus” model of effective PD and begin to build a research-based model of effective science PD. This study also expands our understanding of effective science teacher PD by moving us toward a focus on the substance of PD programs and not just on the form. This study challenges teacher educators to consider new aspects of both the form and substance of their PD and preservice teacher education offerings.

**Summary**

The bottom line for the STeLLA line of research is student learning. Each of the studies in the STeLLA line of research looks at impact on student learning. By documenting significant impact on student learning, the STeLLA line of research is now

- exploring ways to reach additional teacher audiences (e.g., high school science teachers, preservice and first year teachers, K–3 as well as upper elementary teachers),
- exploring ways to adapt the STeLLA program so that it can reach a much wider audience and so that it can be sustained without ongoing support from research dollars, and
• nominating key features of STeLLA professional development that might account for the strong student learning results (moving beyond the consensus model of effective PD).

To develop an understanding of the key features of professional development that make a difference in terms of student learning, we assert that the field of science education research needs to explore new PD research terrain. First, we need to have more lines of research that produce PD programs with well-documented impact on student learning. Then we can look across such studies and ask: What do these programs have in common in terms of both program substance (what teachers are learning) and program form (how are teachers learning)? But this will not be enough to help us develop a strong research-based model of effective science PD. We also need to know about the interactions of features within a program; PD programs like STeLLA have many components that are designed to work together. Which features or which combinations of features predict teacher and student learning? The STeLLA line of research, for example, shows the power of a tight focus on a research-based, conceptual framework; a limited set of science teaching strategies; and science content that is learned largely through analysis of practice. Is the power of the program explained by the presence of a conceptual framework (program substance)? Or is there something about this particular conceptual framework that is important to understand? Or is it the videocases (program form) that make the biggest contribution to the learning results? We need careful thinking about research designs that will help us answer such questions.
References


Reiser, B. (2013). *What professional development strategies are needed for successful implementation of the Next Generation Science Standards?* Princeton, NJ: K-12 Center ETS.


