Research Article

Elementary Teachers' Science Knowledge and Instructional Practices: Impact of an Intervention Focused on English Language Learners

Okhee Lee,[1] Lorena Llosa,[1] Feng Jiang,[2] Alison Haas,[1] Corey O'Connor,[1] and Christopher D. Van Booven[1]

[1]*New York University, 239 Greene Street, Room 620, New York, New York 10003*
[2]*University of Arkansas, Fayetteville, Arkansas*

Received 15 September 2015; Accepted 26 December 2015

Abstract: As part of a three-year curricular and professional development intervention focused on English language learners (ELLs), this study examined the intervention's effect on teachers' science knowledge and instructional practices after one year of implementation. The P-SELL (Promoting Science Among English Language Learners) intervention comprised curriculum materials for students and teachers and teacher professional development workshops during the summer and throughout the school year. Using a cluster randomized controlled trial design, the study involved 103 fifth grade teachers from 33 treatment schools and 116 fifth grade teachers from 33 control schools across three school districts in one state. The teachers completed a researcher-developed science test and a questionnaire about their instructional practices at the beginning and end of the year. Results indicated a positive effect of the P-SELL intervention on teachers' science knowledge and all four measures of instructional practices: teaching for understanding, teaching for inquiry, language development strategies, and home language use. This positive effect could be attributed to consistent implementation of the key features of the intervention (i.e., standards-based, inquiry-oriented, and language-focused) for all students and ELLs in particular through educative curriculum materials and professional development. © 2016 Wiley Periodicals, Inc. J Res Sci Teach 53: 579–597, 2016
Keywords: elementary science; teacher science knowledge; teacher instructional practices; science inquiry; randomized controlled trial

The imperative that all students achieve high academic standards is becoming ever more urgent as a result of several key factors. First, U.S. students have a less than glowing academic track record when compared to their international peers based on the Program for International Student Assessment 2012 results (National Center for Education Statistics [NCES], 2014a) and the Trends in International Mathematics and Science Study 2011 results (NCES, 2013). Second, science achievement gaps persist among demographic subgroups, including the gap between English language learners (ELLs) and non-ELLs (NCES, 2012). Finally, the challenge of high academic standards in science for all students is likely to intensify as the Next Generation Science Standards (NGSS Lead States, 2013), which are being adopted in some states, are both

Contract grant sponsor: National Science Foundation; Contract grant number: DRL 1209309.
Correspondence to: O. Lee; E-mail: olee@nyu.edu
DOI 10.1002/tea.21314
Published online 5 February 2016 in Wiley Online Library (wileyonlinelibrary.com).

academically rigorous and language intensive for all students and ELLs in particular (Lee, Quinn, & Valdés, 2013).

ELLs make up the fastest growing student population in the U.S. According to the 2010 U.S. Census (U.S. Census Bureau, 2012), 21% of children 5–17 years old spoke a language other than English at home. During the 2011–2012 school year, students with limited English proficiency (the term used by the federal government), or ELLs, constituted 9% of public school students or an estimated 4.4 million students (NCES, 2014b). While the ELL student population is rapidly growing, many elementary teachers feel unprepared to teach science to all students, especially ELLs (Banilower et al., 2013).

This study draws from a large research and development project to scale up P-SELL (Promoting Science Among English Language Learners; Lee & Llosa, 2011–2015), a curricular and professional development intervention aimed at enhancing teachers' science knowledge and instructional practices to improve science achievement of all students, especially ELLs. P-SELL consists of a science curriculum for fifth grade students and teachers and professional development workshops across three demographically diverse and geographically disparate school districts in one state. This comprehensive, stand-alone, year-long fifth grade science curriculum is (i) aligned with state science standards; (ii) based on science inquiry both as a goal of science learning and as a means of developing scientific understanding; and (iii) focused on providing guidance and scaffolding for English language development in science instruction. Then, professional development workshops focus on enabling teachers to fully realize the intentions of the curriculum. Thus, P-SELL provides equitable learning opportunities for ELLs in science, while also conceptualizing essential elements of effective science education for all students in the context of high-stakes science assessment and accountability policy.

In the current study, we examined P-SELL's impact on elementary teachers' science knowledge and instructional practices after the first year of implementation. Teachers' science knowledge was assessed using a researcher-developed science test. Teachers' instructional practices were measured using a questionnaire that addressed four domains: teaching practices for understanding, teaching practices for inquiry, language development strategies, and home language use. Specifically, the study examined the following research question: What was the impact of the intervention on teachers' science knowledge and instructional practices after one year of implementation?

Literature Review

Three areas of the literature are discussed: teachers' science knowledge and instructional practices with ELLs, professional development for science instruction with ELLs, and interventions aimed at integrating science learning and language development for ELLs.

Teachers' Science Knowledge and Instructional Practices With ELLs

To promote student learning and achievement, teachers need to know subject matter content. Studies have shown that students working with elementary teachers with higher levels of science knowledge have higher science achievement (Diamond, Maerten-Rivera, Rohrer, & Lee, 2014; Heller, Daeler, Wong, Shinohara, & Miratrix, 2012). Elementary teachers' lack of science knowledge has been related to their inability to teach science effectively (Diamond et al., 2014; Heller et al., 2012; Nowicki, Sullivan-Watts, Shim, Young, & Pockalny, 2013).

Scientific understanding involves deep and complex understanding of science concepts, making connections among concepts, and applying concepts in explaining natural phenomena and real world situations (Kennedy, 1998; National Research Council [NRC], 2007). To enable students to develop scientific understanding, teachers should help students to recognize problematic and incomplete information, make reasoned and well-supported arguments, justify solutions based on evidence, and negotiate ideas and construct collective meanings about science. In addition, to foster students' engagement in science inquiry, teachers should engage students in the practices of science as students ask questions about natural phenomena, construct explanations, argue from evidence based on observations or data, and communicate findings using multiple forms of representation (NRC, 2000, 2012). However, elementary school teachers are often inexperienced when it comes to teaching science to promote students' understanding and engagement in inquiry, and this lack of experience acts as a barrier for teaching science in this manner (Blanchard, Southerland, & Granger, 2009; Harris & Rooks, 2010; Wilson, Taylor, Kowalski, & Carlson, 2010).

Science has a rich language base; thus, the teaching of science and the teaching of language are integrally related (Lee et al., 2013; Rosebery & Warren, 2008). The growing literature indicates four domains of strategies to integrate content and language for ELLs across subject areas including science (see Lee & Buxton, 2013 for detailed descriptions). First, effective teachers incorporate second language (ESOL) and experientially-oriented strategies, including hands-on and purposeful activities, realia, and multiple examples of language in various contexts. Second, effective teachers facilitate ELLs' participation in classroom discourse to help the students understand academic content. Effective teachers are aware of and adaptive to variation in their students' levels of English proficiency and use multiple modes of representation (gestural, oral, pictorial, graphic, and textual). Third, effective teachers focus on students' home language as an instructional support (Goldenberg, 2013; Swanson, Bianchini, & Lee, 2014). Finally, effective teachers capitalize on students' "funds of knowledge" (González, Moll, & Amanti, 2005) by incorporating students' cultural artifacts and community resources in ways that are both academically meaningful and culturally relevant.

Professional Development for Science Instruction With ELLs

The literature indicates core and structural features of effective professional development (Desimone, 2009; Garet, Porter, Desimone, Birman, & Yoon, 2001; Penuel, Fishman, Yamaguchi, & Gallagher, 2007). Core features of effective professional development include (i) focus on content knowledge and how students learn that content; (ii) opportunities for teachers to engage in active learning; and (iii) coherence with other activities for teacher learning and development. In addition, structural features of effective professional development include (i) sufficient duration in terms of the number of contact hours and the duration throughout the year and (ii) collective participation of teachers from the same school, department, or grade level.

In addition, effective professional development allows teachers to fully realize the intentions of the curriculum and utilize the curriculum as scaffolds to promote teacher learning in addition to student learning (Remillard, 2005). Davis and Krajcik (2005) proposed the notion of "educative curriculum materials" that promote changes in teachers' knowledge and practices to make instructional decisions in specific instances and facilitate teachers' development of more general knowledge that can be applied in new situations (Davis et al., 2014; Drake, Land, & Tyminski, 2014).

Finally, effective professional development for science teachers of ELLs in particular enables the teachers to view ELLs' development of literacy and science knowledge as inextricably linked; indeed, contextualized learning in science can afford students an authentic environment, in which to build academic language (Buxton & Lee, 2014; Janzen, 2008; Lee, 2005). Furthermore, ELLs' prior experiences with language (e.g., home language) and culture (e.g., "funds of knowledge", see González et al., 2005) can be utilized as intellectual resources (Warren, Ballenger, Ogonowski, Rosebery, & Hudicourt-Barnes, 2001).

Interventions Aimed at Integrating Science Learning and Language Development for ELLs

Educational interventions that aim to integrate the teaching of science inquiry with the teaching of language development for ELLs have yielded positive results in changing teachers' science knowledge and instructional practices (Bravo, Mosqueda, Solís, & Stoddard, 2014; Hart & Lee, 2003; Lee & Maerten-Rivera, 2012; Shanahan & Shea, 2012). Shanahan and Shea (2012) examined the impact of one professional development model that explicitly embedded language learning strategies into science inquiry lessons. Specifically, the model "focused on how science lessons can include multiple and expanded opportunities for all students to produce oral language" (p. 412). The professional development involved 68 K-2 self-selected teachers from a low-performing school district. Teachers participated in a one-week summer workshop and monthly workshops. To examine the impact of the model, 21 out of the 68 teachers were randomly selected for observations and 6 of the 21 were selected for interviews. The results indicated significant increases in the use of student-talk strategies and increases in teacher confidence in implementing language and science integration.

Bravo et al. (2014) investigated whether a program for preservice elementary teachers that emphasized issues of language and culture in science instruction had an impact on teachers' beliefs and practices. The study involved 65 preservice teachers who participated in the intervention and 45 preservice teachers in the control group. The intervention included participation in a science teaching methods course and a practicum with an experienced teacher who was implementing the program's model in two culturally and linguistically diverse school districts. Using a questionnaire and classroom observations, the researchers found that preservice teachers in the treatment group had stronger beliefs about the effectiveness of the practices promoted by the intervention than preservice teachers in the control group. In their practicum, preservice teachers in the treatment group were observed using questions that elicited higher-order thinking, providing scaffolds, and including language and literacy in their science instruction more frequently than preservice teachers in the control group.

Hart and Lee (2003) examined the impact of an intervention that integrated inquiry-based science, English language and literacy, and students' home language and culture on teachers' beliefs and practices in science and literacy integration. The study involved 53 third and fourth grade teachers from six elementary schools serving diverse student groups. Teachers participated in four full-day workshops throughout the year and were provided with curricular materials for two instructional units per grade. After the first year of participation, the teachers came to place greater emphasis on the importance of reading and writing in science instruction, express a broader and more integrated conceptualization of literacy in science, and provide more effective linguistic scaffolding to enhance scientific understanding.

Lee and Maerten-Rivera (2012) examined change in teachers' knowledge and practices while they participated in a later iteration of the curricular and professional development intervention used in Hart and Lee (2003). The study involved 198 third, fourth, and fifth grade teachers from six urban elementary schools that served predominantly Spanish or Haitian Creole speaking students with low socioeconomic status and that performed poorly on high-stakes state assessments. The intervention consisted of a series of curriculum units that constituted the entire science curriculum from grades 3 to 5 and teacher workshops during the summer and school year. A series of longitudinal multilevel models were used to examine change in teachers' knowledge and practices over time. The results indicated improvements in teachers' knowledge and practices in teaching science to ELLs over the course of the intervention.

Although all of the studies reviewed have found approaches to integrating content and language instruction to be promising, most did not include a control group. The current study used a randomized controlled trial design with a treatment and a control group, thus allowing for a direct investigation of the impact of the intervention on teachers' science knowledge and instructional practices. In addition, the study involved a large-scale implementation of a year-long, stand-alone, comprehensive curriculum for fifth grade in 33 schools across three demographically diverse and geographically disparate school districts in one state. Given the rigorous methodological approach and the scale of the study (33 schools in the treatment group and 33 schools in the control group), this study overcomes the limitations of prior research and thus can address the extent to which interventions for ELLs have an impact on teachers' science knowledge and instructional practices.

The Intervention

The P-SELL intervention comprised (i) curriculum materials for students and teachers and (ii) teacher professional development workshops during the summer and throughout the school year.

Curriculum

Based on the literature on science instruction for ELLs described above and our previous research (Lee, Lewis, Adamson, Maerten-Rivera, & Secada, 2008), we developed a comprehensive, stand-alone, year-long science curriculum for fifth grade.[1] Teachers were provided with complete class sets of curriculum materials (i.e., consumable student books, teacher guide, science supplies, and supplementary materials on the project website). The curriculum highlights three key features described below.

First, the P-SELL curriculum is closely aligned with the state's fifth grade science standards.[2] Each chapter addresses a big idea in nature of science, physical science, Earth and space science, or life science, with the relevant science content standards and benchmarks identified at the outset of each chapter. The focal science content standard(s) and benchmark(s) are also specified for reading passages, writing sections, and inquiry activities.

Second, the P-SELL curriculum is based on an inquiry-oriented approach (NRC, 2000, 2012), particularly taking into account elementary school teachers' lack of knowledge and experience with science inquiry (Blanchard et al., 2009; Wilson et al., 2010). Inquiry serves not only as an end of science learning in itself, but also as a means to build and extend students' knowledge of the science concepts and big ideas targeted by the state science standards. The student book is organized such that teacher-directed instruction increasingly gives way to more student-directed learning; that is, earlier chapters provide more structure, while later chapters promote student initiative and exploration. At the completion of inquiry activities, students are encouraged to design their own extension inquiry activities and apply key science concepts to everyday events or phenomena in home and community contexts.

Finally, the P-SELL curriculum is language-focused as it provides guidance and scaffolding for English language development of ELLs (Goldenberg, 2013; González et al., 2005; Lee & Buxton, 2013). The key science terms for every chapter are provided in the three languages predominantly spoken by the students in the school districts participating in the study—English, Spanish, and Haitian Creole. Then, the chapter introduces key science concepts by relating them to students' prior knowledge or experiences in their home and community contexts, as well as their knowledge from previous chapters. The curriculum represents science content in multiple modes, from textual and graphic (e.g., through extensive writing in the student book) to oral and aural (e.g., through small and whole group discussions). At the end of each chapter, a review of the key science concepts is provided in the form of an expository text, with translations into Spanish and Haitian Creole available on the project website. Additional language development activities and a Spanish translation of the curriculum are available on the project website.

The teacher guide is designed to assist teachers with curriculum implementation based on the notion of educative curriculum materials (Davis et al., 2014; Davis & Krajcik, 2005; Drake et al., 2014). The teacher guide begins with a front matter section that explains key aspects of the curriculum, including (i) how students' mastery of the state science standards is promoted; (ii) the key role of inquiry as a means to develop student understanding of big ideas and key science concepts; (iii) the process of progressively scaffolding students toward self-initiated inquiry; and (iv) strategies for engaging all students, especially ELLs, in rigorous language and literacy development. For each chapter, following the science inquiry activities, the teacher guide provides science background information and explanations for the questions under investigation and related natural phenomena. These explanations address recurring learning difficulties that students frequently encounter during each of the inquiry activities. In addition, the teacher guide provides instructional strategies tailored to the specific science content for each chapter. For example, it provides suggestions on setting up and implementing hands-on activities, paired with insights on how to avoid or resolve common problems. The teacher guide also provides suggestions for different levels of guidance and scaffolding by using additional activities for students who need support for content mastery as well as enrichment activities for students who need challenge beyond content mastery. Finally, it includes language development strategies and web-based resources to promote science learning for all students, particularly ELLs.

The fifth grade science teachers in the treatment schools implemented the intervention curriculum, whereas the fifth grade science teachers in the control schools implemented the district-adopted textbooks ("business as usual"). District A used "Interactive Science" by Pearson. The fifth grade textbook consists of four units: science, engineering, and technology; life science; earth science; and physical science. The textbook was supplemented with district-created labs and activities. District B used "National Geographic Science." The fifth grade materials included Big Ideas Student Books focusing on earth science, life science, and physical science that served as the base textbook. "National Geographic Science" also includes a separate science inquiry/writing book and access to an interactive student website. Teachers can supplement their instruction with a science-inquiry materials kit, but no supply replenishments were provided. District C used "Science Fusion" by Houghton Mifflin Hartcourt. The fifth grade textbook consists of 10 units in four areas: the nature of science, earth and space science, physical science, and life science. "Science Fusion," similar to "National Geographic Science," offers an Interactive Student Edition that students can write in and access to an interactive student website with virtual labs. "Science Fusion" includes hands-on activities with science supplies provided, but with no replenishments.

The P-SELL curriculum differs from all three textbook series with regard to three key features. First, the P-SELL curriculum is closely aligned to the state science standards and stays within the content limits of these standards, whereas the district-adopted textbooks go beyond the content limits, leaving teachers to decide curriculum coverage. Second, the P-SELL curriculum incorporates fully-developed inquiry activities that are essential and closely connected to science concepts addressed in chapters. The P-SELL curriculum also provides enough science supplies for students themselves to conduct the hands-on inquiry activities, and consumable supplies are replenished each year. In contrast, the district-adopted textbooks have fewer and less fully developed hands-on inquiry activities that are optional and not closely connected to science concepts addressed in lessons or chapters. They often provide supplies for teacher demonstrations, and do not replenish consumable supplies. Finally, the P-SELL curriculum focuses on English language development for ELLs and offers extensive supplementary materials on the website including language development activities for ELLs of lower proficiency levels, whereas the district-adopted textbooks do not explicitly consider English language development for ELLs.

Teacher Workshops

During the first year of the intervention, teachers in the treatment group attended four or five full-day workshops during the summer and throughout the school year and one full-day year-end meeting for data collection, feedback, and planning for the following year. The timelines varied among the three school districts to meet each district's guidelines. District A offered a three-day summer workshop shortly before the school year started, two full-day workshops during the school year, and the year-end meeting. District B also offered a three-day summer workshop shortly before the school year started, two full-day workshops during the year, and the year-end meeting. District C offered a one-day summer workshop shortly after the school year started, three full-day workshops during the year, and the year-end meeting. Teachers received stipends for attending the summer workshops, and schools received payments for substitute teachers during the school year.

During the first year, the focus of the workshops was on classroom implementation of the P-SELL curriculum. Teachers became familiarized with the three main features of the P-SELL curriculum: (i) alignment to state science standards; (ii) hands-on science inquiry and understanding; and (iii) language development strategies with ELLs, as described below.

The workshops highlighted the state science standards and content limits of these standards, so that the intervention was coherent with expectations for science instruction and demands for high-stakes assessments in science. For example, the standards for each chapter were communicated explicitly with teachers at the introduction of the chapter during the workshop, as the standards similarly appear at the top of the chapter in the teacher guide.

A primary goal of the workshops was to promote inquiry-based science. During the workshops, teachers performed every hands-on inquiry activity in the curriculum, as their students would do throughout the year. They discussed aspects that might go wrong in the classroom, possible errors they wanted to avoid, and intentional "errors" they wanted their students to experience as learning opportunities. By engaging fully in the inquiry activities, teachers had opportunities to use science content in explaining the results of inquiry activities, applying the science content to new situations, and asking questions for further investigations. Through this process, teachers experienced firsthand what classroom discourse might look like during inquiry activities and discussed how they could facilitate classroom discourse for all students and ELLs in particular before, during, and after the inquiry activities.

Strategies for English language development were introduced as they were embedded in hands-on science inquiry and understanding. Teachers discussed how to utilize second language pedagogies and strategies typical of contextualized experiential approaches, classroom discourse strategies, students' home language as an instructional support (Goldenberg, 2013), and "funds of knowledge" (González et al., 2005). For example, teachers explored science terms in students' home language and cognates between English and the home language. Teachers also explored how bilingual students can assist less English proficient students in their home language.

As the workshops continued over the school year, teachers implemented the P-SELL curriculum in their teaching and reflected on their practices. The workshops promoted collective participation of all fifth grade science teachers within each school and each school district. Teachers were given time for collaborative planning to develop common goals, share resources, and exchange ideas and experiences arising from the common context of the intervention. The networks generated during the workshops gave teachers opportunities to build social capital that otherwise would not be afforded them.

Method

Research Setting

The study was implemented in three school districts in a southeastern state. According to the NCES (2006), District A located in the northeastern part of the state was designated as urban, District B located in the southwestern part of the state was designated as urban/suburban, and District C located in the central part of the state was designated as urban/suburban. The three districts encompassed a wide range of racial, ethnic, socioeconomic, and linguistic diversity. During the 2012–2013 school year, District A was 45% Black, 40% White non-Hispanic, 8% Hispanic; 52% participated in the free or reduced price lunch (FRL) program; and 3% of the student population was designated as limited English proficient (LEP, the federal term). District B was 51% White non-Hispanic; 28% Black, 15% Hispanic; 52% participated in FRL; and 8% was LEP. District C was 34% Hispanic, 30% Black, 28% White non-Hispanic; 60% participated in FRL; and 14% was LEP.

Research Design

During the 2012–2013 school year, District A had 103 elementary schools, District B 44 elementary schools, and District C 125 elementary schools. A cluster randomized controlled trial was conducted. Within each of the three school districts, 22 schools were randomly selected to participate, yielding a total of 66 participating schools. Within each district, half of the selected schools were randomly assigned to the treatment group and half to the control group, yielding a total of 33 treatment schools and 33 control schools across the three districts. As shown in Table 1, the random assignment procedures yielded equivalent treatment and control groups in terms of school-level student demographic variables.

Teacher Participants

All fifth grade science teachers in the 66 schools participated in the study. During the first year of the intervention (2012–2013), the study involved 258 teachers. Of the 258 teachers, 30 were in their first year of teaching science. Since the pre-questionnaire administered in the beginning of the school year asked teachers to report on their science instructional practices during the prior year, there was no baseline data available for these 30 teachers and thus they were not included in the analyses. Among the remaining 228 teachers, three did not complete the questionnaire in the fall and six did not complete the questionnaire in the spring because they were no longer working

Table 1
School characteristics

	Treatment ($n=33$)		Control ($n=33$)		Overall ($n=66$)				
Variable	*M*	SD	*M*	SD	*M*	SD	Diff	*t*	*p*
% of students receiving FRL	71.2	21.7	71.5	20.3	71.3	20.9	0.4	0.07	0.94
% of white students	34.1	21.5	33.7	22.2	33.9	21.7	−0.4	−0.07	0.94
% of black students	31.6	28.6	34.9	26.9	33.2	27.6	3.4	0.50	0.62
% of Asian students	4.3	5.0	2.9	2.5	3.6	4.0	−1.5	−1.50	0.14
% of Hispanic students	26.9	21.6	25.4	18.2	26.1	19.8	−1.5	−0.30	0.77
% of ESE students	11.9	3.4	12.3	3.8	12.1	3.6	0.4	0.44	0.66
% of ELL students	11.3	12.9	8.5	11.1	9.9	12.1	−2.8	−0.93	0.36

at the schools. The final sample included 219 teachers, 103 in the treatment group and 116 in the control group.

Tables 2 and 3 show the demographic and professional backgrounds of the teachers in the sample. The majority of the teachers were female, White, non-Hispanic, and native speakers of English. Most of the teachers had a bachelor's degree (63%) and met the state's requirements for ESOL preparation via endorsement through the school district (61%). There were no differences between the treatment and control groups in terms of years of teaching, years of teaching science, number of science courses taken, or number of science methods courses taken.

At the end of the school year, teachers reported time for professional development in science and ESOL during the school year. The teachers in the treatment group were asked to exclude the five full day workshops that were part of the intervention. As shown in Table 4, 37% of control teachers and 47% of treatment teachers reported that they did not attend any science workshops during the year. In addition, 72% of control teachers reported that they did not attend any ESOL workshops compared to 54% of treatment teachers. In general, professional development opportunities in teaching science and working with ELLs were limited for the teachers in the study (aside from the 5–6 full day workshops that were part of the intervention for the treatment group).

Even though professional development opportunities were limited, teachers in both groups taught science regularly and extensively (as shown in Table 5), which reflects the fact that

Table 2
Comparison of treatment and control teacher background variables (dichotomous and categorical variables)

Variable	Control ($n = 116$) (%)	Treatment ($n = 103$) (%)
Demographic background		
Gender		
Female	85	80
Male	15	20
Ethnicity		
White, non-Hispanic	74	76
Black, non-Hispanic	13	12
Hispanic or Latino	12	10
Asian	1	2
Native language*		
English	96	98
Spanish	8	8
Other fluent language*		
English	4	2
Spanish	2	5
Professional background		
Highest degree		
Bachelor's degree	61	65
Master's degree or higher	39	35
ESOL training*		
Met ESOL requirement through college coursework	22	26
Met ESOL requirement through school district (META)	64	57
Completed bachelor's or master's degree in ESOL	6	7
Other	9	16
None	4	1

* Teachers could select more than one response.

Table 3
Comparison of treatment and control teacher background variables (continuous variables)

Variable	Control ($n = 116$)		Treatment ($n = 103$)		Overall ($n = 219$)				
	M	SD	*M*	SD	*M*	*SD*	Diff	*t*	*p*
Professional background									
Years of teaching	13.2	9.23	12.1	9.1	12.7	9.1	1.1	0.85	0.40
Years of teaching science	10.9	7.81	9.7	6.5	10.3	7.2	1.2	1.29	0.20
Science courses	2.6	2.61	3.2	3.1	2.9	2.9	−0.7	−1.67	0.10
Science methods courses	1.6	1.66	1.5	1.1	1.5	1.4	0.1	0.35	0.72

high-stakes science assessment counted toward school accountability in fifth grade: 58% of the teachers spent 151–300 minutes teaching science each week and another 16% taught science more than 300 minutes each week. Both the total minutes of science instruction per week and the average length of science class were comparable between the treatment and control groups.

Instruments

Data were collected using two instruments, a science test and a questionnaire, that were completed by teachers in both the treatment and control groups at the beginning and end of the year.

Table 4
Frequency and time for professional development

Variable	Control (%)	Treatment (%)
Science workshops—how often		
Never	37	47
Once	30	27
Twice	16	10
3–4 times	11	9
More than 4 times	5	8
Science workshops—how many hours		
N/A	37	47
1–6 hours	43	33
7–15 hours	6	7
16–35 hours	11	9
More than 35 hours	3	5
ESOL workshops—how often		
Never	72	54
Once	20	21
Twice	2	6
3–4 times	4	7
More than 4 times	3	12
ESOL workshops—how many hours		
N/A	72	55
1–6 hours	20	24
7–15 hours	4	5
16–35 hours	2	8
More than 35 hours	3	8

Table 5
Time for science instruction

Variable	Control (%)	Treatment (%)
Science instruction per week		
Less than 60 minutes	4	6
60–150 minutes	25	17
151–300 minutes	58	58
More than 300 minutes	13	19
Length of science class		
30–45 minutes	38	26
46–60 minutes	42	54
61–75 minutes	16	17
76–90 minutes	4	2

Science Test. Two equated forms of a science test were developed to measure teachers' science knowledge (see Maerten-Rivera, Huggins, Adamson, Lee, & Llosa 2015). The two forms are composed of public release items from the National Assessment of Educational Progress (NAEP), Trends in International Mathematics and Science Study (TIMSS), and selected state science assessments. Each form has a spread of low, medium, and high difficulty items across the four "bodies of knowledge" including Earth and space science, life science, physical science, and nature of science. To capture increase of teachers' science knowledge over time and to avoid a ceiling effect, the selected items measure topics covered in fifth grade but were designed for eighth or tenth grade students.

Each of the two test forms has 30 multiple-choice items and three short response items. Of the 30 multiple-choice items, six remain the same across the two forms for linking purposes. The three short response items also remain the same on both forms of the test. Each multiple-choice item was worth one point, one open-ended item was worth two points, and the other two open-ended items were worth four points each. The maximum score possible was 40 points. Specifications for form 1 of the teacher science test are available in the online Supplementary Materials.

For the scoring of the three short response items, we adopted the NAEP scoring rubric that corresponded to each item. A team of raters participated in a two-hour training session prior to scoring responses to each item. All the tests were independently scored by two raters. Disagreements were resolved by a third round of scoring and group consensus, if needed. The inter-rater agreement for all three items in the pretest and posttest was excellent (weighted Kappa above 0.75). The internal consistency, Cronbach's α, of the test scores for the pretest (form 1) and posttest (form 2) were 0.78 and 0.79, respectively.

The test development process included content review, initial tryouts, pilot testing, and item analyses in a manner that was consistent with accepted standards of assessment design (AERA/APA/NCME, 1999). The initial version of the test was piloted with 175 elementary school teachers, including 100 teachers who took form 1 and 75 teachers who took form 2. Elementary teachers from the three school districts in the study did not participate in the pilot testing. Details about the development and validation of the test are reported in Maerten-Rivera et al. (2015).

Questionnaire. The questionnaire used in this study was a refined version of the questionnaire used in our previous research (see Lee & Maerten-Rivera, 2012, for details). It asked for information about teachers' demographic and professional background (highest degree, years of teaching, number of science courses taken, number of science methods courses taken, years of teaching science, and ESOL training). Then, it measured teachers' self-reported practices in teaching science in four domains: (i) teaching practices to promote students' scientific understanding; (ii) teaching practices to promote students' scientific inquiry; (iii) teaching practices to support language development; and (iv) home language use with ELLs. These four constructs were grounded in the relevant literature, described earlier. Each of the four scales consisted of three to seven items. Each individual item asked teachers to rate the frequency or intensity of reform-oriented practices using a 4-point rating system (1 = never or almost never; 2 = some lessons; 3 = most lessons; 4 = every lesson). The four scales including all of the items are presented in the online Supplementary Materials.

The score for each scale was computed using the average of the responses to the items that composed the scale. Use of the average item response, as opposed to the summated score, ensured that missing responses would not lead to a systematic negative bias of the scale scores. A scale score was computed only for those respondents who had valid responses for at least 75% of the items in the scale. If someone answered fewer than 75% of a scale's items, the respondent's scale score was set to be missing and omitted from analyses. The reliability of the scale scores was estimated using Cronbach's α. Internal reliability estimates for all of the scales were above the acceptable level of 0.70. The estimates ranged from 0.76 to 0.84 in the pre-questionnaires and from 0.73 to 0.83 in the post-questionnaires (see Table 6 below).

Data Collection

In the treatment group, the teacher science test and questionnaire were administered prior to the start of the intervention during the first teacher workshop and at the completion of the first-year implementation during the year-end workshop. In the control group, the science test and the questionnaire were administered at the school sites in the beginning of the school year and at a combination of school district buildings and school sites at the end of the school year.

Data Analysis

To examine the impact of the intervention on teachers' science knowledge (as measured by the science test) and instructional practices (as measured by four questionnaire scales), we used multilevel modeling (also known as hierarchical linear modeling; Raudenbush & Bryk, 2002). Results of fitting unconditional two-level models indicated that across all five outcomes the school-level intraclass correlations (ICC) ranged from 0.09 to 0.13 (see Table 6). The data had a two-level structure whereby teachers (level-1)

Table 6
Descriptive statistics for science test scores

				Control			Treatment			
	ICC	Time	α	*n*	*M*	SD	*n*	*M*	SD	*t*	*p*
Teacher knowledge	0.13	Pre	0.80	116	28.0	5.48	103	28.2	5.87	0.27	0.787
		Post	0.83	116	29.3	5.78	103	30.9	5.56		
Teaching practices for understanding	0.13	Pre	0.79	113	2.80	0.51	98	2.79	0.54	−0.25	0.802
		Post	0.75	116	2.87	0.44	103	3.10	0.44		
Teaching practices for inquiry	0.09	Pre	0.79	112	2.28	0.46	98	2.34	0.49	0.83	0.409
		Post	0.81	116	2.46	0.44	103	2.66	0.50		
Language development strategy	0.10	Pre	0.76	110	2.85	0.48	98	2.89	0.53	0.69	0.49
		Post	0.73	115	2.93	0.45	102	3.19	0.45		
Home language use	0.13	Pre	0.84	74	1.61	0.68	73	1.97	0.84	2.71	0.008
		Post	0.83	72	1.79	0.76	66	2.33	0.99		

were nested in schools (level-2). Therefore, two-level models were used with two dummy coded variables for district included at level-2. To describe these models, we denote the score on the teacher posttest score for the *i*th teacher of the *k*th school by $POST_{ik}$. The general form of the multilevel model is given by:

Level-1 model: $POST_{ik} = \pi_{0k} + \pi_1(PRE_{ik}) + e_{ik}$

Level-2 model: $\pi_{0k} = \gamma_{00} + \gamma_{01}(TRT_k) + \gamma_{02}(Dist1_k) + \gamma_{03}(Dist2_k) + \gamma_{04}(Pre_k) + r_{0k}$

where PRE_{ik} corresponds to the teacher pretest score centered at the school mean, TRT_k corresponds to school condition (treatment vs. control), $Dist1_k$ and $Dist2_k$ correspond to two dummy coded variables representing districts, Pre_k corresponds to the pretest school mean centered at the district mean, and errors are denoted by e_{ik} and r_{0k}. The overall treatment effect is represented by the coefficient γ_{001}. Separate multilevel models were estimated for the science test and each of the four instructional practices scales.

Since we were testing five related outcomes, we applied the Benjamini–Hochberg correction for adjusting *p*-values (Benjamini & Hochberg, 1995).[3] Finally, we calculated effect sizes by dividing the unstandardized regression coefficient for the treatment effect (i.e., an estimate of the difference between treatment and control at posttest) by the pooled standard deviation of the posttest.

Results

Table 6 shows descriptive statistics for the teacher science test and the four questionnaire scales at pre and post, as well as *t*-test results comparing means between the treatment and control groups at pre.

Impact of the Intervention on Teachers' Science Knowledge

As shown in Table 6, at the start of the intervention, the average score on the teacher science test was 28.0 (out of 40 maximum points) for control teachers and 28.2 for treatment teachers, indicating that on average teachers got about 70% of the maximum points. There was no statistically significant difference between treatment and control teachers in terms of their science knowledge as measured by the science test ($t = 0.27$, $df = 217$, $p = 0.787$).

The results of the multilevel analyses, shown in Table 7, indicate that the intervention had a positive impact on teachers' science knowledge: the treatment effect was statistically significant ($p = 0.005$) and the effect size was 0.24, a small effect size.

Impact of the Intervention on Teachers' Instructional Practices

As shown in Table 6, there were no statistically significant differences between treatment and control teachers in terms of the frequency, in which they reported teaching practices for understanding ($t = -0.25$, $df = 206$, $p = 0.802$), teaching practices for inquiry ($t = 0.83$, $df = 206$, $p = 0.409$), or use of language development strategies ($t = 0.69$, $df = 199$, $p = 0.490$). Teachers in the treatment group, however, reported more frequent use of home language than teachers in the control group ($t = 2.71$, $df = 143$, $p = 0.008$).

The results of the multilevel analyses indicate that the intervention had a statistically significant impact on all four questionnaire scales related to teachers' instructional practices. Treatment teachers reported more frequent use of teaching for understanding practices ($p < 0.001$), teaching for inquiry practices ($p = 0.002$), language development strategies ($p < 0.001$), and home language use ($p = 0.020$) (see Table 7). The effect sizes for these four scales were moderate, ranging from 0.41 to 0.56.

Table 7
Treatment effect on teachers' science knowledge and instructional practices

Term	Teacher Knowledge	Teaching Practices for Understanding	Teaching Practices for Inquiry	Language Development Strategies	Home Language Use
N					
School	65	63	63	64	46
Teacher	219	211	210	206	111
Fixed Effects					
Intercept	6.57**	1.69***	1.37***	1.71***	1.17***
Treatment	1.39**	0.23***	0.19**	0.25***	0.45*
Pretest (school average)	0.03	0.03	0.18	0.15	0.08
Pretest (teacher level)	0.79***	0.39***	0.28**	0.28**	0.26*
District 1	0.18	0.00	0.04	−0.05	−0.02
District 2	0.07	0.02	0.04	0.05	0.21
Random effects					
School	0.13	0.01	0.00	0.00	0.05
Teacher	11.61***	0.14***	0.18***	0.15***	0.60***
Treatment effect size	0.24	0.52	0.41	0.56	0.52

*$p < 0.05$.
**$p < 0.01$.
***$p < 0.001$.

Discussion

The study examined whether P-SELL, a comprehensive, stand-alone, year-long curricular and professional development intervention focused on ELLs and implemented large-scale across three school districts in one state had an impact on fifth grade teachers' science knowledge and instructional practices. At the start of the intervention, there were no significant differences between the treatment and control groups in teachers' science knowledge or reported instructional practices, except for home language use.

After one year of implementation, the P-SELL intervention had a positive effect on teachers' science knowledge and all four scales of practices in teaching science with ELLs, with teachers who implemented P-SELL scoring higher on the science test and reporting more frequent use of instructional practices than teachers in the control group. These impacts were found, despite the fact that the P-SELL intervention was implemented large-scale in 33 schools under routine conditions, and despite the fact that teachers in the control group taught science as regularly and extensively as teachers in the treatment group. Both groups were motivated to teach science because science was a tested subject in fifth grade in the state where the study took place.

The positive impacts could be attributed to the key features of the P-SELL intervention (i.e., standards-based, inquiry-oriented, and language-focused). First, the curriculum materials (student books, teacher guide, and science supplies), which were standards-based, inquiry-oriented, and language-focused, were designed to be educative (Davis & Krajcik, 2005; Davis et al., 2014; Drake et al., 2014). Second, the teacher workshops, which followed the core and structural features of effective professional development (Desimone, 2009; Garet et al., 2001), emphasized the standards-based, inquiry-oriented, and language-focused approach of the intervention. For example, to enable teachers to promote mastery of state science standards through science inquiry for all students, especially ELLs, the teacher workshops offered opportunities for teachers to engage in science inquiry themselves while at the same time gaining knowledge of the state science standards. The workshops also highlighted language development strategies for ELLs during science inquiry. Thus, the focus on the key features of the P-SELL intervention implemented consistently through the curriculum materials and teacher workshops might have resulted in the positive impacts.

Although previous studies focused on interventions aiming to integrate science and language education have generally found positive results for teacher learning and change, strong claims could not be made about the benefits of these types of interventions because of their methodological limitations (e.g., no control groups). The current study overcomes these limitations and thus can make significant contributions to the literature. First, the current study employed a randomized controlled trial design, in which schools were randomly assigned to a treatment group and a control group. Thus, change in teachers' science knowledge and instructional practices can be directly attributed to the intervention. Second, the study included a teacher sample from schools that were randomly selected from all schools in each school district. And, unlike previous studies that were conducted in one location, the current study was conducted in 66 schools across three demographically diverse school districts in one state, thus enhancing the generalizability of the findings. Furthermore, the intervention involved all fifth grade teachers in the participating schools, rather than a self-selected group of volunteer teachers. Finally, the study used a teacher science test (Maerten-Rivera et al., 2015) and a questionnaire (Lee & Maerten-Rivera, 2012) that were carefully designed and validated, thus addressing the lack of measures that is often cited as an obstacle to studying the impact of professional development interventions on teacher outcomes (Desimone, 2009; Wayne, Yoon, Zhu, Cronen, & Garet, 2008). Our measures, along with the description about the development of the science test (Maerten-Rivera et al., 2015) and the teacher questionnaire (Lee & Maerten-Rivera, 2012), make an important contribution to this literature. Overall, the current study responds to the call for the use of experimental designs in studies of teacher professional development (Wayne et al., 2008) and contributes to the literature on scale-up of educational interventions (McDonald, Keesler, Kauffman & Schneider, 2006).

One limitation of the current study is that it relied on teachers' reported instructional practices. Given the large-scale nature of the project, observations of teachers' instructional practices were not possible. For classroom observations to yield reliable information about instructional practices, a teacher should be observed multiple times (Desimone, 2009). Since our sample included 258 teachers, it was not possible for us to conduct meaningful classroom observations given the scope and budget of the project. Even observing one randomly selected teacher per school (66) was not a viable option. Given the great variability between teachers within a school, the observations gathered from one teacher could not be generalized to other teachers in the school, and thus the classroom observations would have been of limited use. Instead, we relied on a teacher questionnaire. According to Desimone (2009), surveys can provide "valid and reliable data on the amount of time that teachers spend on specific practices occurring during a set time period—up to a year" (p. 190), the type of data we collected using the teacher questionnaire. Trade-offs involved in the use of different types of measures is an ongoing discussion in the literature on professional development interventions (Desimone, 2009; Wayne et al., 2008).

The results of our study suggest areas for further research. In our project, additional time points will be used to extend this study as teachers continue their participation in the P-SELL intervention over the three-year period of the project. We will be able to examine how teachers' science knowledge and instructional practices change over time. For example, in examining the effect of a 3-year intervention on teachers' perceptions of instructional practices, Supovitz, Mayer, and Kahle (2000) and Lee and Maerten-Rivera (2012) found growth during the first year of participation in the intervention, which was sustained in the subsequent two years, that is, "short-term growth and long-term stability" (Supovitz et al., 2000, p. 342). In their 4-year longitudinal study of Cognitively Guided Instruction, Fennema et al. (1996) found that 18 of the 21 participating teachers showed positive changes in their instructional practices and beliefs by the final year. These studies did not include control groups and thus could not directly attribute growth or lack thereof to the intervention.

The study examined the effect of P-SELL, a large-scale intervention across varied educational settings in the context of science accountability policy. The results indicate that a standards-based, inquiry-oriented, and language-focused science intervention delivered through educative curriculum materials and high-quality professional development can enable elementary teachers to improve their science knowledge and instructional practices. Design, implementation, and testing of science educational interventions for teachers and students will become increasingly important as some states adopt the NGSS that are both academically rigorous and language intensive for all students and ELLs in particular (Lee et al., 2013).

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Notes

[1] Sample lessons are available upon request.

[2] It is noted that the curriculum was not aligned with the Next Generation Science Standards (NGSS). The curriculum development preceded the publication of the NGSS and the state, in which the study took place did not adopt the NGSS during the period when our intervention was implemented.

[3] To apply the Benjamini–Hochberg method, we followed the steps outlined in the What Works Clearinghouse Procedures and Standards Handbook, Version 3.0, pages G1-G5. In this method, the *p*-values are first sorted and ranked. The smallest value gets rank 1, the second rank 2, and the largest gets rank N. Then, each *p*-value is multiplied by its assigned rank and divided by N to give the adjusted *p*-values.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). Standards for educational and psychological testing. Washington, DC: APA.

Banilower, E. R., Smith, P. S., Weiss, I. R., Malzahn, K. A., Campbell, K. M., & Weis, A. M. (2013). Report of the 2012 national survey of science and mathematics education. Chapel Hill, NC: Horizon Research, Inc.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B Methodological, 57(1), 289–300.

Blanchard, M. R., Southerland, S. A., & Granger, E. M. (2009). No silver bullet for inquiry: Making sense of teacher change following an inquiry-based research experience for teachers. Science Education, 93(2), 322–360.

Bravo, M. A., Mosqueda, E., Solis, J. L., & Stoddard, T. (2014). Possibilities and limits of integrating science and diversity education in preservice elementary teacher preparation. Journal of Science in Teacher Education, 25, 601–609.

Buxton, C. A., & Lee, O. (2014). English language learners in science education. In N. G. Lederman, & S. K. Abell (Eds.), Handbook of research in science education, 2nd ed. (pp. 204–222). Mahwah, NJ: Erlbaum.

Davis, E., & Krajcik, J. (2005). Designing educative curriculum materials to promote teacher learning. Educational Researcher, 34(3), 3–14.

Davis, E. A., Palincsar, A. S., Arias, A., Bismack, A., Marulis, L., & Iwashyna, S. (2014). Designing educative curriculum materials: A theoretically and empirically driven process. Harvard Educational Review, 84(1), 24–52.

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. Educational Researcher, 38(3), 181–199.

Diamond, B. S., Maerten-Rivera, J., Rohrer, R. E., & Lee, O. (2014). Effectiveness of curricular and professional development intervention on elementary teachers' science content knowledge and student achievement outcomes: Year 1 results. Journal of Research in Science Teaching, 51, 635–658.

Drake, C., Land, T. J., & Tyminski, A. M. (2014). Using educative curriculum materials to support the development of prospective teachers' knowledge. Educational Researcher, 43(3), 154–162.

Fennema, E., Carpenter, T. P., Franke, M. L., Levi, L., Jacobs, V. R., & Empson, S. B. (1996). A longitudinal study of learning to use children's thinking in mathematics instruction. Journal for Research in Mathematics Education, 27(4), 403–434.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. American Educational Research Journal, 38, 915–945.

Goldenberg, C. (2013). Unlocking the research on English learners: What we know—and don't yet know—about effective instruction. American Educator, 37(2), 4–11.

González, N., Moll, L. C., & Amanti, C. (2005). Funds of knowledge: Theorizing practices in households, communities, and classrooms. Mahwah, NJ: Erlbaum.

Harris, C. J., & Rooks, D. L. (2010). Managing inquiry-based science: Challenges in enacting complex science instruction in elementary and middle school classrooms. Journal of Science Teacher Education, 21(2), 227–240.

Hart, J., & Lee, O. (2003). Teacher professional development to improve science and literacy achievement of English language learners. Bilingual Research Journal, 27(3), 475–501.

Heller, J. I., Daeler, K. R., Wong, N., Shinohara, M., & Miratrix, L. W. (2012). Differential effects of three professional development models on teacher knowledge and student achievement in elementary science. Journal of Research in Science Teaching, 49(3), 333–362.

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

Kennedy, M. M. (1998). Education reform and subject matter knowledge. Journal of Research and Science Teaching, 35, 249–263.

Lee, O. (2005). Science education and English language learners: Synthesis and research agenda. Review of Educational Research, 75, 491–530.

Lee, O., & Buxton, C. A. (2013). Integrating science learning and English language development for English language learners. Theory Into Practice, 52(1), 36–42.

Lee, O., Lewis, S., Adamson, K., Maerten-Rivera, J., & Secada, W. G. (2008). Urban elementary school teachers' knowledge and practices in teaching science to English language learners. Science Education, 92(4), 733–758.

Lee, O., & Llosa, L. (2011–2015). Promoting science among English language learners (P-SELL) scale-up. National Science Foundation, Discovery Research K-12.

Lee, O., & Maerten-Rivera, J. (2012). Teacher change in elementary science instruction with English language learners: Results of a multi-year professional development intervention across multiple grades. Teachers College Record, 114(8), 1–44.

Lee, O., Quinn, H., & Valdés, G. (2013). Science and language for English language learners in relation to next generation science standards and with implications for common core state standards for English language arts and mathematics. Educational Researcher, 42(4), 223–233.

Maerten-Rivera, J., Huggins, A. C., Adamson, K., Lee, O., & Llosa, L. (2015). Development and validation of a measure for elementary teachers' science content knowledge in two multiyear teacher professional development intervention projects. Journal of Research in Science Teaching, 52(3), 371–396.

McDonald, S-K., Keesler, V. A., Kauffman, N. J., & Schneider, B. (2006). Scaling-up exemplary interventions. Educational Researcher, 35(3), 15–24.

National Center for Education Statistics. (2006). Data local education agency locale code file: School year 2005–06. Washington, DC: U.S. Department of Education, Institute of Education Sciences. Retrieved from http://nces.ed.gov/ccd/pdf/pau051agen.pdf

National Center for Education Statistics. (2012). Science 2011: National assessment of educational progress at grade 8 (NCES 2012-465). Washington, DC: U.S. Department of Education.

National Center for Education Statistics. (2013). Highlights from TIMSS 2011: Mathematics and science achievement of U.S. fourth- and eighth-grade students in an international context (NCES 2013-009). Washington, DC: U.S. Department of Education.

National Center for Education Statistics. (2014a). Performance of U.S. 15-year-old students in mathematics, science, and reading literacy in an international context: First look at PIS 2012 (NCES 2014-024). Washington, DC: U.S. Department of Education.

National Center for Education Statistics. (2014b). The condition of education 2014 (NCES 2014-083). Washington, DC: U.S. Department of Education.

National Research Council. (2000). Inquiry and the national science education standards: A guide for teaching and learning Washington, DC: National Academy Press.

National Research Council. (2007). Taking science to school: Learning and teaching science in grades K-8. Washington, DC: National Academy Press.

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

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

Nowicki, B. L., Sullivan-Watts, B., Shim, M. K., Young, B., & Pockalny, R. (2013). Factors influencing science content accuracy in elementary inquiry science lessons. Research in Science Education, 43(3), 1135–1154.

Penuel, W. R., Fishman, B., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. American Educational Research Journal, 44(4), 921–958.

Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical linear models: Applications and data analysis methods. 2nd ed. Thousand Oaks, CA: Sage.

Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. Review of Educational Research, 75(2), 211–246.

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

Shanahan, T., & Shea, L. (2012). Incorporating English language teaching through science for K-12 teachers. Journal of Science Teacher Education, 23, 407–428.

Supovitz, J. A., Mayer, D. P., & Kahle, J. B. (2000). Promoting inquiry-based instructional practice: The longitudinal impact of professional development in the context of systemic reform. Educational Policy, 14(3), 331–356.

Swanson, L. H., Bianchini, J. A., & Lee, J. S. (2014). Engaging in argument and communicating information: A case study of English language learners and their science teacher in an urban high school. Journal of Research in Science Teaching, 51(1), 31–64.

U.S. Census Bureau. (2012). Statistical abstract of the United States, 2012. Washington, DC: Government Printing Office. Retrieved from http://www.census.gov/commpendia/statab/cats/education.html

Warren, B., Ballenger, C., Ogonowski, M., Rosebery, A., & Hudicourt-Barnes, J. (2001). Rethinking diversity in learning science: The logic of everyday language. Journal of Research in Science Teaching, 38(5), 529–552.

Wayne, A. J., Yoon, K. S., Zhu, P., Cronen, S., & Garet, M. S. (2008). Experimenting with teacher professional development: Motives and methods. Educational Researcher, 37(8), 469–479.

What Works Clearinghouse Procedures and Standards Handbook, Version 3.0, pages G1–G5. Retrieved from https://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v3_0_standards_handbook.pdf

Wilson, C. D., Taylor, J. A., Kowalski, S. M., & Carlson, J. (2010). The relative effects and equity of inquiry-based and commonplace science teaching on students' knowledge, reasoning, and argumentation. Journal of Research in Science Teaching, 47(3), 276–301.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.