Journal of Educational Change (2019) 20:221–248
https://doi.org/10.1007/s10833-019-09338-2

Coping with constraints: Longitudinal case studies of early elementary science instruction after professional development

Judith Haymore Sandholtz[1] · **Cathy Ringstaff**[2] · **Bryan Matlen**[2]

Published online: 14 March 2019

Abstract
Professional development holds significant potential in promoting science education, but that potential is undermined if instructional changes are not sustained. For professional development, sustainability refers to the continuation of outcomes over an extended period of time after the program ends and is an issue across international contexts. This longitudinal research, part of a larger research project funded by NSF, investigated sustainability of early elementary science instruction after teacher participation in a 3-year professional development program and the factors that influenced teachers' decisions about instructional time and strategies. The research used a case-study approach to obtain an in-depth understanding of why instructional shifts occurred over a period of 7-years. The primary data sources were teacher surveys, self-efficacy assessments, and interviews. The findings highlight how school level factors changed over time and affected teachers' science instruction in both positive and negative ways. The research illustrates the connection between contextual constraints and science instruction and holds implications for sustaining meaningful instructional changes after professional development ends.

Keywords Elementary education · Longitudinal research · Teacher professional development · School context · Science instruction · Sustainability

✉ Judith Haymore Sandholtz
judith.sandholtz@uci.edu

Cathy Ringstaff
cringst@wested.org

Bryan Matlen
bmatlen@wested.org

1 School of Education, University of California, Irvine, 3200 Education Building, Irvine, CA 92697-5500, USA

2 WestEd, 400 Seaport Court, Suite 222, Redwood City, CA 94063-2767, USA

Introduction

> The legacy that it [the professional development] has left for me is to keep teaching science as much as you can even though your site administrator or your district may not push for it anymore.

Miranda, a second-grade teacher quoted above, drew upon personal passion to continue teaching science 4-years after the professional development program ended. But other teachers, such as Robert, gradually succumbed to contextual constraints and started to reduce science instruction before finally reverting to short, textbook lessons. Miranda and Robert were two participants in a 3-year state-funded professional development program that provided science assistance for teachers in rural schools in the United States. Earlier research about the impacts of the professional development reported significant changes in the participating teachers' science content knowledge, self-efficacy related to science teaching, instructional time devoted to science, and instructional practices in science after just 1-year of the professional development; these changes continued during the second and third years of the program (Sandholtz and Ringstaff 2013a, b). Participating teachers described shifting from teaching science whenever they could "fit it in" to establishing set time slots for science in their instructional plans. In interviews, they explained the value in using more hands-on, experimental, and investigative strategies. But 4-years after the program ended, the persistence of these outcomes varied widely across teachers and schools.

Although professional development holds significant potential in promoting science education in elementary schools, that potential is undermined if changes are not sustained. For professional development, sustainability refers to the continuation of outcomes over an extended period of time after the program ends and is an issue across international contexts. In the field of education, reform efforts often focus on increasing the scale of implementation rather than sustaining changes (McLaughlin and Mitra 2001). Although the assumption that effects of professional development programs will be maintained over time is not self-evident, research on sustainable impact is generally lacking (Zehetmeier and Krainer 2011).

Professional development outcomes may be difficult to sustain for a range of reasons. Changing teachers' pedagogical practices initially is not a quick process, but rather one that often extends over years of professional development (Darling-Hammond et al. 2009; Hawley and Valli 1999). During the professional development, teachers tend to benefit from supports generated by the programs. To sustain newly-adopted pedagogical changes after the professional development ends may require ongoing forms of support that are not readily apparent or available. These supports may include, for example, periodic refresher sessions, collaboration with other teachers, materials for teaching science, webinars, or electronic support. A report from the National Academy of Sciences (2015) in the U.S. concludes that science teachers' development should be perceived not only as long term but also as contextualized. Contextual factors at multiple levels affect the extent to which teachers adopt instructional strategies learned in professional development as well as the extent to which any instructional changes persist over time. Gaikhorst et al. (2017) contend that the effectiveness of professional development depends on program characteristics, but the sustainability of program effects is related more to characteristics of school organizations.

Research on teacher professional development has focused primarily on the important features of effective programs and outcomes at the end of the programs. There is relatively little research on long-term effects of professional development, perhaps due to the challenges of collecting longitudinal data from teachers. However, given the investments of time and money in professional development, more research about sustainability is needed. To understand if and how programs can lead to sustained outcomes, researchers need to examine factors that both promote and hinder sustainability after the programs end.

This study addresses the need for research that investigates sustainability of professional development outcomes. Specifically, this study examined longitudinal changes in early elementary science instruction after teacher participation in a 3-year professional development program and the factors that influenced teachers' decisions about instructional time and strategies. The research, part of a larger NSF-funded project, used a case-study approach to obtain an in-depth understanding of why instructional shifts occurred over 7-years. The primary research questions were: To what extent did teachers continue to teach science and use inquiry-based instructional strategies after the professional development ended? What factors influenced science instruction for these teachers?

Conceptual framework

Operating from the premise that all children can learn science and benefit from scientific reasoning, science educators have focused on preparing teachers to teach science to students in diverse settings (Settlage et al. 2017). However, the Committee on Strengthening Science Education (National Academy of Sciences 2015) in the U.S. recognizes "a gap between what science teaching and learning could be and the reality of current practices" (p. 2). A key strategy for addressing this gap is professional development. Researchers report that professional development not only increases teachers' preparedness in science content and pedagogy (Banilower et al. 2007; Bowes and Banilower 2004; Duschl et al. 2007; Gess-Newsome 2001; Penuel et al. 2008; Supovitz and Turner 2000), but it also promotes changes in teaching practices (National Staff Development Council 2001; Rotermund et al. 2017; Sparks 2002; Stigler and Hiebert 1999). Desimone (2009) proposes an operational theory of how professional development leads to changes in teachers' instructional practices. The core theory of action includes four steps. First, teachers participate in effective professional development. Second, their participation increases their knowledge and skills or changes their attitudes and beliefs. Third, given their new knowledge and skills (or attitudes and beliefs), teachers adapt their instructional practices through changes in content, pedagogy, or both. Fourth, the changes in instructional practices promote student learning. In this model, context functions as a key mediating influence and includes, for example, student characteristics, teacher characteristics, curriculum, school leadership, policies at multiple levels, and classroom, school, and district environments. This study focuses on the third step of Desimone's model, changes in instructional practice, and devotes particular attention to context.

Although studies primarily focus on characteristics of effective professional development and outcomes, researchers acknowledge the influence of contextual factors on the extent to which teachers design high quality curriculum and strategies (Penuel and Gallagher 2009) and implement new instructional practices in their classrooms (Guskey and Sparks 2002; National Academy of Sciences 2015; Penuel et al. 2007). Changes in science instruction depend not only on individual teachers' abilities and attitudes, but also on contextual support and factors beyond their control (Banilower et al. 2007; Johnson 2006; Sherry 2002; Haney et al. 2002). Consequently, researchers report that professional development providers "need to consider not only teachers' own contexts but also the program's demands on teachers and how those demands can be met within their contexts" (Penuel et al. 2007, p. 952).

Administrative support is an important contextual factor in promoting science education, particularly in elementary schools. Research on the Local Systemic Change program, which included 42 projects, reported that support from the school principal predicted how much instructional time teachers spent on science and to what extent they used investigative strategies (Banilower et al. 2007). Adequate administrative support contributes not only to positive changes in teachers' practices stemming from professional development (Guskey and Sparks 2002) but also teachers' continued use of reform-based strategies (National Academy of Sciences 2015; Steele 2001). Administrators also have a role in resource availability, another contextual factor that affects teachers' implementation of investigative science lessons (Johnson 2006; Sullivan-Watts et al. 2013; Wyner 2013). In addition to resources such as lesson ideas and reference materials, classroom materials such as science kits and supplies for experiments foster the use of investigative, inquiry-based strategies (Appleton and Kindt 1999; Steele 2001; Sullivan-Watts et al. 2013).

Collaboration among teachers and collegial support also contribute to implementation of classroom teaching strategies learned in professional development (Franke et al. 2001; Johnson 2006; Pea 2012). Support from colleagues affects elementary teachers' confidence in teaching science as well as their classroom instruction (Appleton and Kindt 1999). However, opportunities for collaboration that teachers experience during professional development programs may not be available to the same extent at their school sites and thereby affect whether or not teachers continue to use the focal instructional strategies.

Accountability pressures stemming from reform efforts, standardized testing, and district and state policies influence science instruction in various ways. In elementary schools with high accountability pressure, typically related to student achievement in math and language arts, teachers spend significantly less time using hands-on and laboratory teaching methods in science (Hayes and Trexler 2016). Moreover, the extent to which standards for science education are implemented by teachers is influenced by policies and decisions at the district level (Spillane and Callahan 2000). Standards-based reform has different, and often unintended, consequences on science education practices depending on school contexts (Settlage and Meadows 2002).

Schools located in rural settings encounter heightened challenges stemming from contextual factors. Compared to urban and suburban schools, rural schools are usually smaller in terms of student populations and are located in sparsely populated but widespread areas (Monk 2007). In addition, student populations typically include higher numbers of minority students and English language learners (Jimerson 2005) and students living in poverty (Avery 2013; Jimerson 2005; Monk 2007). Rural schools also confront financial constraints, increased transportation costs, and outdated and small facilities (Farmer 2009; Harmon and Smith 2007; Jimerson 2005). These issues make it particularly challenging to attract and retain qualified STEM teachers in rural settings (Avery 2013). Teachers in rural schools typically have less teaching experience than other teachers and more limited backgrounds in science (Arnold et al. 2004). Moreover, providing content-specific professional development is problematic due to tight budgets, large distances between schools, and limited organizational support (Harmon et al. 2007), and rural teachers have less access to other educational opportunities offered by science organizations, colleges, and corporations (Avery 2013). Given the difficult decisions that administrators in rural settings face (Farmer 2009), support for science education in elementary schools may take a lesser priority.

In summary, context is a key mediating influence on changes to teaching practices stemming from professional development. Research on sustainability is lacking, but researchers propose that school characteristics may have more influence than program characteristics on the extent to which professional development outcomes are sustained over time (Gaikhorst et al. 2017). As described above, schools in rural settings face additional challenges related to contextual factors. According to Guskey (2002), the most neglected aspect of professional development is sustaining changes after programs end. This study, which specifically examines whether teachers' changes in instructional practices endured over time and what factors influenced their decisions about science instruction, addresses a need for longitudinal research that investigates sustainability of professional development outcomes. If science is not taught on a consistent basis and if pedagogical practices in science are not sustained over time, improved student learning in science in early elementary classrooms is unlikely.

Methods

Professional development program

This study is part of a larger longitudinal research project examining the sustainability of professional development outcomes. The research project focuses on early elementary teachers who completed a 3-year professional development program that provided science assistance for teachers in rural districts in the U.S. The professional development included three key components: (a) intensive adult-level science content instruction; (b) pedagogical training focused on science instruction and how to connect science to language arts and mathematics; and (c) training and support designed to facilitate teacher collaboration. The program provided teachers with over 100 contact hours each year and included intensive summer institutes, regional meetings, and school site sessions. The authors of this paper were not involved in designing or providing the professional development, but the second author served as an external evaluator.

The content instruction during the summer institutes focused on a different branch of science each year (physical, earth, and life sciences) and was based on topics included in the California state science standards. A team that included a university professor with expertise in science and advanced mathematics, an elementary teacher with expertise in research-based instructional strategies and science inquiry, and an English language learning specialist led the 6-day summer institutes. Following each summer institute, teachers participated in regional meetings as well as sessions at their schools during the academic year.

The pedagogical component, which emphasized scientific inquiry, was intertwined with the adult-level content instruction. Teachers learned science content through research-based instructional strategies that included scientific inquiry as well as hands-on experiments and investigations. The pedagogical component also introduced teachers to a model of instruction called Scaffolded Guided Inquiry (SGI) that they could use back in their classrooms (Vanosdall et al. 2007). The central aim of this guided inquiry process is to promote learning through student investigation. The inquiry-based instruction mirrors scientific methods and engages students in higher-level thinking and science process skills such as making predictions, summarizing knowledge, analyzing data, and evaluating their findings. SGI lessons are meant to take place over several days and include nine steps: (a) display the big idea; (b) gather needed materials; (c) discuss an engaging scenario; (d) identify a focus question; (e) make a prediction; (f) collect data; (g) make claims based on evidence; (h) draw a conclusion; and (i) reflect. As part of the inquiry process, students are supposed to write in science notebooks and record their focus question, prediction, data, claims and evidence, conclusions, and reflections.

The professional development program also helped teachers create inquiry-based science units based on the SGI model through the use of a curriculum-mapping tool. The tool provides a process for documenting curriculum, planning for implementation of standards, and matching assessment with instruction. As part of the pedagogical component, teachers also learned instructional strategies shown to be effective for English Language Learners and approaches for integrating science instruction with mathematics and language arts.

To enhance collaboration among teachers in these rural schools, the program created opportunities for teachers to work together. For example, during summer institutes, teachers worked in teams to develop curriculum maps for various grade levels. With the aim of creating professional learning communities, the program organized teachers into cluster groups according to the geographic proximity of their schools. During each school year, teachers participated in regional "cluster" meetings where they could reunite to discuss implementation, share strategies, and plan events. The program also hosted a website for teachers to communicate, share lesson plans, and access instructional resources.

Participants

The 3-year program included 39 teachers from 16 schools in 16 districts in northern California in the U.S. Half of the districts were one-school districts in which a particular grade level may have only one teacher. Student enrollment prior to the program ranged from 148 to 5087, and student performance on standardized tests indicated low academic achievement. Due to changes in teaching assignments, relocations, death, and attrition, there were 34 participating teachers by the end of the 3-year program. Our longitudinal follow-up research project includes 30 of those teachers, representing 14 schools and 13 districts.

In this study, we used a comparative case study approach with a purposive sample of four teachers. A case study design is particularly well suited to examining how and why contemporary events occur (Yin 2003). Comparative cases facilitate examination of factors that may mediate teachers' professional practice but may not be evident with single cases or cases that do not differ in ways important to the general problem (Patton 2002). Since the aim was to investigate why teachers made varied decisions about science instruction after the professional development ended, the teacher was the unit of analysis. To select teachers, we used survey data from the larger research project to examine teachers' instructional time in science over a 7-year period: 3-years of the professional development and 4-years after the program ended. We plotted teachers' responses to two key questions: (a) number of minutes in a typical science lesson and (b) number of days that science is taught in a typical week. To examine critical cases, which had "strategic importance in relation to the general problem" (Flyvberg 2001, p. 78), we reviewed the plots and selected four teachers who had differing patterns of instructional time in science and who taught in different schools. The patterns included: low instructional time that gradually increased over the years (Miranda); gradual decrease in time before a small rise followed by a sharp decrease (Robert); increases to high instructional time before dropping to moderate amounts (Gabriela); overall high instructional time but yearly variations that included both increases and decreases (Rachel).

Although instructional time was the key variable for selecting teachers, the four teachers are representative of the larger group of K-2 teachers (primarily white females with a range of teaching experience). The student population at their schools ranged from approximately 200–800 and approximately 5–55% English language learners. Their teaching assignments included first and second grades, and the number of other participating teachers at their schools ranged from zero to four. The school contexts are described more specifically in each case study.

Data sources and analysis

Primary data sources for the case studies included teacher surveys, self-efficacy assessments, and interviews. The teacher survey, developed by Horizon Research, Inc (2000) for national studies of science teaching at the elementary level, focused on teachers' opinions about science and science instruction, their preparedness, and their instructional practices. The self-efficacy assessment, the Science Teaching Efficacy Beliefs Instrument (STEBI) developed by Riggs and Enochs (1990) specifically for use with elementary teachers, focused on teachers' beliefs about their effectiveness in teaching science. Teachers, as part of the larger research project, completed the teacher surveys and self-efficacy assessments at seven time points: pre-program, in each of the three program years, and in each of the three post-program years. To examine perceptions of principal support and use of instructional strategies and student activities, we examined a series of questions from the teacher survey (see Table 1). Teacher ratings were converted to numeric values on a five-point scale, and composite scores were calculated. For each case study teacher, we plotted composite scores across the 7-year period. To examine

Table 1 Items from the teacher survey

Category	Survey item
Instructional strategies in science	Introduces content through formal presentations
	Demonstrates a science-related principle or phenomenon
	Teaches science using real-world concepts
	Arranges seating to facilitate student discussion
	Uses open-ended questions
	Requires students to support their claims with evidence
	Encourages students to explain concepts to one another
	Encourages students to consider alternative explanations
	Allows students to work at their own pace
	Helps students see connections between science and other disciplines
	Uses assessment to gauge what students know before or during a unit
	Embeds assessment in regular class activities
	Assigns science homework
	Reads and comments on students' reflections
Student activities during science instruction	Participates in discussions with teacher
	Works in cooperative groups
	Works on solving a real-world problem
	Shares ideas or solves problems in small groups
	Engages in hands-on activities
	Designs or implements their own investigation
	Works on extended science investigations
	Records, represents, and/or analyzes data
	Writes reflections in notebook
	Uses mathematics to solve problems
Principal support for science	Encourages selection of science content and strategies that address individual learning styles
	Encourages implementation of national standards in science education
	Encourages innovative instructional practices
	Provides needed materials and equipment
	Provides time for teachers to meet and share ideas
	Encourages making connections across disciplines
	Acts as a buffer between teachers and external pressures

teachers' self-efficacy over time, we plotted their overall scores on the self-efficacy assessment for the seven administrations.

As part of the larger research project, researchers conducted in-person or telephone interviews with teachers 2, 3, and 4-years after the program ended. Each interview lasted approximately 1 h and was audio-recorded and subsequently transcribed. These semi-structured interviews focused on instructional time in science, confidence in teaching science, content knowledge in science, instructional and curricular choices, integration of science into other subjects, and support and resources for teaching science. The questions probed for changes in the years after the professional development ended and reasons for any reported changes.

Analysis of interview data followed qualitative research procedures such as coding and data displays (Bogdan and Biklen 1998). All verbatim interview transcripts were compiled in an electronic database and coded initially according to a priori codes from the conceptual framework and interview protocols. A second round of coding identified emergent sub-categories; for example, sub-categories in the contextual factors category included curricular demands, administrative support, teacher support, and resources. The software allowed for multiple codes to be assigned to interview excerpts, which aided in searching for and retrieving data across categories. For example, teachers' descriptions of science instruction in their classrooms may have included information related to multiple codes such as "strategies used", "Scaffolded Guided Inquiry", "changes in strategies", and "factors influencing strategies". The code "lack of" could be assigned as a double code to any of the categories. To establish reliability in coding, we used a system of multiple coders for each transcript. All coders received training in the coding system and demonstrated competence in applying the codes on a sample transcript. Two coders then independently coded each interview transcript; the coders achieved inter-rater reliability higher than 85% based on the percentage of matches (Creswell 2005). Any discrepancies between the two were noted, discussed, and resolved. The discrepancies were minor and centered on the additional codes for an excerpt. A third coder subsequently reviewed all coded transcripts for consistency with the system.

For this study, we searched the database to retrieve and organize interview data for the case study teachers, reviewed their full interview transcripts, and created data displays for the teachers. Data analysis focused on explanation building (Yin 2003), specifically on understanding why instructional shifts for each teacher occurred over the 7-year period. Using the plots for each teacher, we examined shifts in reported instructional time, self-efficacy, administrative support, and use of instructional strategies and student activities. We subsequently looked in the interview data for excerpts related to each of these categories and for emerging influences on each teacher's instruction in science. For example, one teacher described in an interview how bus schedules influenced the length of the instructional day—an unexpected contextual factor that was not captured in the survey data. In constructing each case, we focused on the teachers' explanations and rationales for their decisions about science instruction after the professional development ended. After constructing each teacher case, we conducted a cross-case analysis, first examining similarities and differences across teachers. Our aim was to look for patterns and key factors across teachers and settings and to understand how these factors influenced teachers' decisions about science instruction over time. The small sample of teachers limits the generalizability of the findings to other teachers, but the findings may suggest areas for future longitudinal research and potential approaches for sustaining professional development outcomes.

Individual cases

In the following sections, we present each of the four case studies and discuss four aspects of the teachers' science teaching: (a) their teaching assignments; (b) changes in instructional time in science after the professional development ended; (c) the impact of the professional development program; and (d) factors that influenced their decisions about science instruction. After presenting the four individual cases, we propose and describe five key findings from our cross-case analysis that relate to sustainability of professional development outcomes. These findings address: (a) teachers' desire to teach science in relationship to constraints; (b) changes in teachers' use of inquiry-based instructional strategies in science; (c) positive and negative influences on instructional time for science; (d) shifts in school level support; and (e) teachers' personal commitment to teach science.

Miranda

Teaching assignment

Throughout the 3-year professional development and 4-year follow-up study, Miranda taught second grade at a small elementary school that enrolled approximately 400 students. The student population consisted of approximately 76% Hispanic and 21% White students. Approximately 47% of the students in the school were English language learners, and over 90% qualified for free and reduced lunch. The percentage of students qualifying for free and reduced price lunch as part of the U.S. National School Lunch Program is often used as an indicator of the concentration of poverty within a school. However, the percentage may overestimate the actual number of students living in poverty.

Change in instructional time

As shown in Fig. 1, Miranda reported little or no science teaching before the professional development program but, over the years, she gradually increased the time allotted to a typical science lesson from 10–20 min to 30–40 min. The number of days vacillated between 1 and 2 days a week.

Impact of professional development

Miranda, who has "always been fascinated with science", described the professional development program as providing her "that booster shot of love of science that

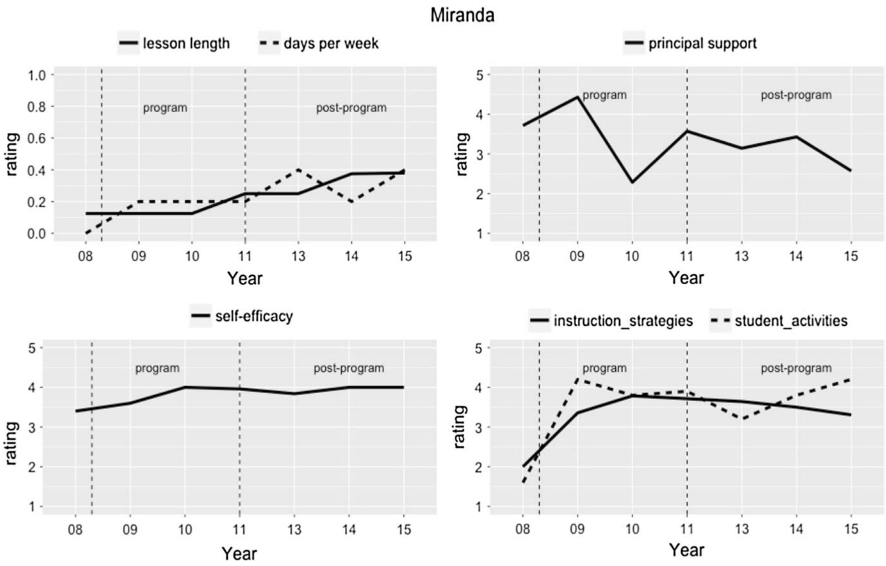

Fig. 1 Change in teacher ratings related to (top left) instructional time: minutes per day in typical lesson (10 or fewer = 0; 11–20 = .13; 21–30 = .25; 31–40 = .38; 41–50 = .5; 51–60 = .625) and days of science instruction per week (none = 0; one = .2; two = .4; three = .6; four = .8; five = 1); (top right) principal support (1 = strongly disagree to 5 = strongly agree); (bottom left) self-efficacy (1 = low to 5 = high); and (bottom right) instructional strategies and student activities (1 = never to 5 = all or almost all)

I already had". Given her passion for the subject, she primarily "wanted to know how I can do it better". She wanted more information about tools, strategies, and materials so she could teach science in "a way that is going to make the students as excited and passionate about it as I am". Miranda was unsure that she "had all the level of knowledge that I needed" but she was never intimidated about teaching science. She credited the professional development program with improving her content knowledge, particularly in physical science, but she most appreciated learning about inquiry-based strategies for use across all three branches of science. For example, she described being uncertain how to teach physical science in a "meaningful way to students rather than just using the textbook which is pretty boring", but coming away from the professional development with "some really good ideas" and activities, such as race car experiments for teaching force and motion, that she took "straight to the classroom".

Miranda identified "learning the SGI model" as a key benefit of participating in the 3-year professional development program. She viewed the model not only as a valuable strategy for teaching content but also as "an excellent way" to teach "that whole scientific methods that they need to know in the upper grades". She also highlighted the way in which the hands-on, investigative strategies increased student engagement in science: "They love science. If they can do those hands-on activities, they are so engaged they want to go on and on with whatever activity… They get so energized, so excited, it is amazing". She also found the strategy helpful with English learners because "It is so hands-on and every lesson starts out with understanding the vocabulary of what you are working with… and we draw a picture of it as part of our inventory chart". Unlike other teachers who ended up eliminating some steps of the model due to time constraints, Miranda reported doing all of the steps but extending SGI lessons over more days or weeks:

> Sometimes you have to draw it out over time. It's still really meaningful for them because you're covering so many standards. All the writing that is involved and the prediction and the summarizing, and then you go back to your focus question, and then you have time to reflect and wonder and come up with more questions.

As shown in Fig. 1, Miranda's self-efficacy in teaching science increased during the professional development program and stayed stable over the subsequent years. After the program ended, she had only one additional opportunity for science-related professional development. The focus was lesson study, and she and a second-grade teacher from a different school collaborated on developing, teaching, and refining a lesson based on rocks and minerals.

Factors that affected science instruction

Miranda consistently identified time as the primary factor that made it difficult to teach science. Although finding instructional time for science is a common issue for elementary teachers, Miranda faced additional time constraints due to the school's ongoing designation as a program improvement site. Schools that do not meet statewide proficiency goals (related to student performance and state standards) for two consecutive years are designated program improvements schools by the state and are subject to improvement and corrective action measures. In an effort to improve student test scores, the school principal contracted with a program to focus on mathematics and language arts standards. In addition to required workshops, meetings, and walk-throughs, Miranda felt "a lot of stress to implement the strategies in this [new] program" and was trying "to figure out how to fit science in" during the afternoon when she often had to "finish what I didn't get to in language arts and math". She described the press of preparing students for benchmark exams in mathematics and language arts and the challenge of "still having time for science". 2-years after the professional development ended, in an interview scheduled in October, Miranda indicated that she had not "even started with science" yet that school year. She thought that "time and consistency" were particularly important when using the SGI model: "You have to be consistent and work on this model. You have to use this model often enough so that they get used to forming predictions and conclusions… You just have to do it enough times". However, due to the pacing schedule and all of the practice tests in language arts and math, she increasingly felt like "I don't have enough time to teach science".

The emphasis on student performance in math and language arts was compounded by turnover in school administrators. Although Miranda remained at the same school, the school principal changed four times during the 7-year period of this research, and Fig. 1 shows how she reported different levels of support by the principals. Miranda indicated that each principal had differing priorities. For example, whereas a previous principal supported her efforts to hold a family science night and helped secure teacher buy-in so that "every teacher at every grade level was doing some kind of science project in their classroom", a subsequent principal wanted to hold a family literacy night instead. She reported that science instruction was "never really discussed or talked about" with any of the principals with whom she worked, and that there was no means for securing needed materials. Miranda also reported little collegial support for science instruction. She was the only person from her school, not just her grade level, who opted to participate in the 3-year professional development program. She stated there was little or no support for science instruction "coming from the district or the site, so it is just a matter of what I decide to do as a teacher or as a grade level. I haven't really had much support from my grade level over any of these years… So whatever I did in science I did on my own".

Despite the contextual challenges and lack of support, Miranda seemed determined to continue to teach science and to incorporate investigative strategies. As shown in Fig. 1, her use of inquiry-based and student-centered instructional strategies increased significantly after the first year of the professional development and changed only slightly after it ended. She expressed concern that many teachers facing time constraints "are reduced to almost the worst strategy to teach science—reading boring stuff out of a book". Even though the professional development had ended, she resolved that, "The legacy that it has left for me is to keep teaching science as much as you can even though your site administrator or your district may not push for it anymore".

Summary

After the professional development, Miranda gradually increased her time teaching science, but struggled to fit science into the instructional day over the years. Requirements related to the school's designation as a program improvement site, frequent changes in administrators and their priorities, and little support from colleagues made it ever more challenging to teach inquiry-based science on a consistent basis. Rather than abandon science instruction or investigative strategies, she pushed to teach science as much as possible within the contextual constraints.

Rachel

Teaching assignment

Throughout the 7-years of this research, Rachel taught first grade at the same elementary school. The school had a population of just under 600 students that consisted primarily of White (77%) and Hispanic (12%) students. Approximately 54% of the students qualified for free and reduced lunch.

Change in instructional time

As shown in Fig. 2, Rachel taught science for approximately 30–40 min twice a week before the professional development started. Over the next 6-years, the number

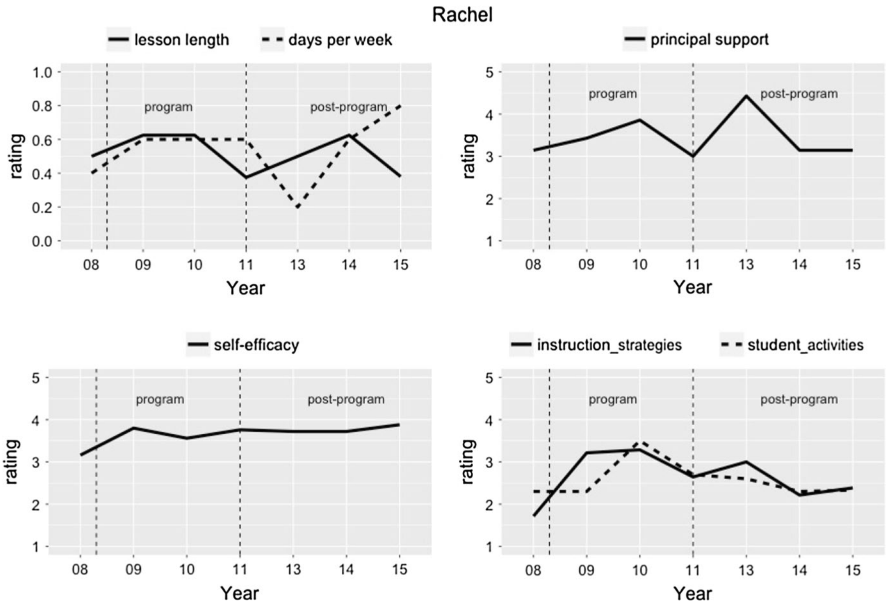

Fig. 2 Change in teacher ratings related to (top left) instructional time: minutes per day in typical lesson (10 or fewer = 0; 11–20 = .13; 21–30 = .25; 31–40 = .38; 41–50 = .5; 51–60 = .625) and days of science instruction per week (none = 0; one = .2; two = .4; three = .6; four = .8; five = 1); (top right) principal support (1 = strongly disagree to 5 = strongly agree); (bottom left) self-efficacy (1 = low to 5 = high); and (bottom right) instructional strategies and student activities (1 = never to 5 = all or almost all)

of days that she taught science varied from 2 to 4 days per week, and the number of minutes in a typical lesson ranged from 30 to 60 min. In some years, Rachel taught science for longer time periods but fewer days a week. For example, at the end of the 3-year program, she described teaching science "every Tuesday and Thursday afternoon" for approximately 90 min as well as "here and there" on other days.

Impact of professional development

Rachel's self-efficacy in teaching science increased after the program's first year and then stayed stable over the subsequent years (see Fig. 2). Given her limited science coursework as an undergraduate, Rachel viewed increased content knowledge as an important benefit of her participation in the professional development. In addition to gaining a "better background" in the three branches of science addressed in the program, Rachel pointed out the value of learning more about the standards and how to incorporate inquiry-based strategies into her science teaching. She credited the professional development with not only building her content and pedagogical knowledge but also fostering her interest in teaching science:

> I have not gone back to the textbook. I feel like [the professional development program] has made science much more interesting and so much more fun to teach.

> My love for science has changed… I'm always searching for new ideas and new wonderful things to add to my lessons and I think that has changed. I never sit and say, 'Okay, these are our lessons and this is what we're doing' and then do it year after year.

She said that teaching science involved active learning for both her and her students. In the years after the professional development ended, Rachel chose to complete two one-week science-related professional development programs—one on oceans and another on forests—that increased her own knowledge and resulted in lesson plans that could be used in her classroom.

Factors that affected science instruction

For Rachel, the factors that most influenced her decisions about science instruction were the standards and students' interest. The standards provided the foundation for organizing the science curriculum, and students' interest reinforced her choice of instructional strategies. Noting that the students' love of science made her "want to do more", she tried to incorporate more hands-on activities to keep students actively engaged:

> I don't think reading a textbook is going to get them excited. I think getting into the activity and maybe reading the directions of the experiment. They have to be engaged… When kids are interested, they are going to go search for more.

To add more science investigations, Rachel started "Fun Fridays", which she described as "some quick Friday science things that they can do within a short time. It has nothing to do with the full science unit I am doing at the time. That is so they can get in there and do a little experiment".

Over the years, Rachel's use of inquiry-based, student-centered strategies declined slightly (See Fig. 2). In interviews, she described continuing to use the SGI model promoted in the professional development but adapting it to take less time and to fit the students' writing skills. She thought it was important for the students to "make predictions, observe, and chart what they're observing, and then write about what the outcome was". But she dropped some steps such as the engaging scenario and reduced the total amount of journal writing, focusing "more on the prediction with their writing and what they found", because "the writing sometimes for first graders is too much". At the time of Common Core State Standards implementation, a national initiative to establish consistent educational standards across the states in the U.S., Rachel expressed some concerns about having spent "the last 5-years building these programs" only to have "complete changes with the science curriculum coming".

When asked what would help her continue to teach science, Rachel identified funding and time for collaboration:

> It would be nice to have more money. It's always good to have time to collaborate with peers and build a stronger program. Schools never give us

> enough time for that, to just collaborate with other teachers… and be able to build a stronger program that gets better every year instead of putting together a unit and using it over and over again.

Rachel was the only first grade teacher at her school who participated in the professional development program. After the program ended, she no longer collaborated with participants from other districts or even participants in other grade levels at her school. However, she reported discussing science with other first-grade teachers at some grade level meetings. Although the school did not have extra funding for science, Rachel benefited from other school-level factors. For example, the parent-teacher association contributed resources for a school science night and parents volunteered their time in classrooms. She stated that the parent support was particularly helpful when conducting science labs and pointed out that she had at least one parent for "the groundhog activity" and "four parents and myself when we dissected owl pellets". Although Rachel reported varying levels of support with changes in school administrators (see Fig. 2), the principals typically allowed teachers flexibility in terms of curriculum and instruction. The teachers were not required to adhere to daily schedules or strict pacing guides in mathematics and language arts and could opt to integrate science with other subjects as they deemed appropriate. As she stated, "We are really on our own. Nobody tells us that we can't teach it [science] or when we have to teach what… We can do it, but we have to find the money".

Perhaps the most striking school-level factor that facilitated science instruction in the early elementary grades was time in the instructional day. The school enrolls students in grades kindergarten through eighth grade but has only one school bus run. Consequently, the instructional day extends from 8:15 a.m. to 2:50 p.m. for all grade levels rather than only the upper grades. In interviews, Rachel acknowledged the challenge that most first grade teachers face trying to fit science into a shorter instructional day. She noted that while the longer instructional day allowed her to teach more science, it did not provide the time needed to "be prepped and ready to go the next day". Despite the time required for preparation, Rachel viewed science as "the good part of the day". As she summarized, "To be honest, the science part of teaching makes me happy. I like it, and I think the kids like it".

Summary

After the professional development ended, Rachel's instructional time in science both increased and decreased over the years but remained higher than other teachers. Her students' interest motivated her use of inquiry-based strategies, and she found ways to adapt investigations to take less time and better fit her students' skills. Rachel benefited from contextual factors such as flexibility in daily schedules, parent volunteers, and a bus schedule that increased the length of the instructional day.

Robert

Teaching assignment

Robert taught second grade at the same elementary school throughout the 7-year period of the research. The school had a student population of approximately 860 students, including 62% Hispanic students and 53% English language learners. Over 90% of the students qualified for free and reduced lunch.

Change in instructional time

As shown in Fig. 3, the number of days that Robert taught science increased from 1 day pre-program to 3 days after the program's first year and stayed consistent thereafter. In contrast, the time allotted to a typical science lesson did not increase from pre-program but rather stayed about 30 min until dropping to 11–20 min 4-years after the professional development ended.

Impact of professional development

Having completed a biology major in college, Robert stated that he "gained some content knowledge [from participating in the professional development] but probably benefited more from the methodology than the content". The program provided opportunities for "collaborating with other teachers on ideas" and "trying out and

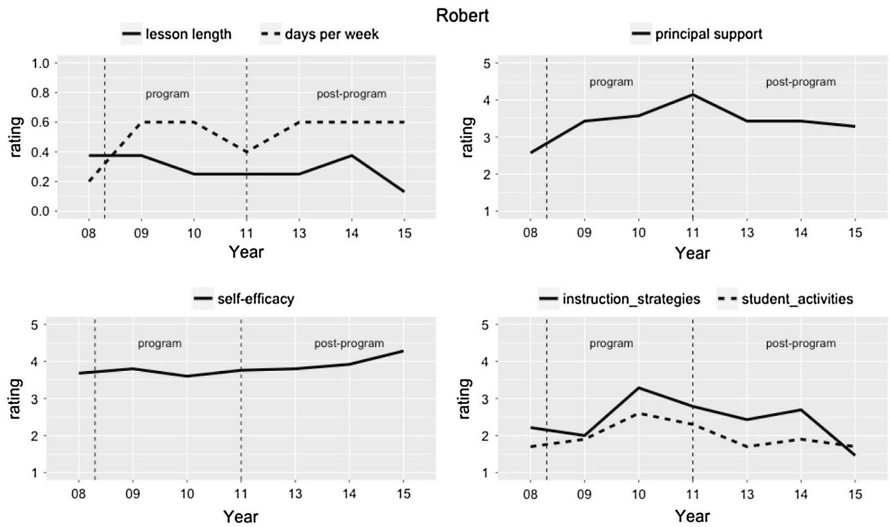

Fig. 3 Change in teacher ratings related to (top left) instructional time: minutes per day in typical lesson (10 or fewer = 0; 11–20 = .13; 21–30 = .25; 31–40 = .38; 41–50 = .5; 51–60 = .625) and days of science instruction per week (none = 0; one = .2; two = .4; three = .6; four = .8; five = 1); (top right) principal support (1 = strongly disagree to 5 = strongly agree); (bottom left) self-efficacy (1 = low to 5 = high); and (bottom right) instructional strategies and student activities (1 = never to 5 = all or almost all)

seeing what works with the kids". For example, before the program, he was unsure how to carry out hands-on, inquiry-based activities with students: "I wasn't sure what to do. I wasn't really clear on what the standards were. You interact with other people and think, 'I think the kids could handle that. I could do that in the classroom'". Once he began using the strategies, he recognized their value for students: "I wasn't doing hands-on [before the professional development] and now I do. And they do remember and get it better when they experience it". He found it helpful to have students "do the hand-on [activities] first" and "to predict what they think will happen" before reading about it in the textbook. For instance, he brought in fruit flies for students to follow the life cycle, and the class tried growing corn in different ways such as "taking away the light, taking away the water, putting it in a jar, cutting the roots off". He reported that his students' excitement about science motivated him and boosted his confidence in teaching it.

Factors that affected science instruction

Robert's survey results show a rise in overall self-efficacy in teaching science yet a decline in his use of inquiry-based and student-centered instructional strategies (see Fig. 3). In interviews after the program ended, Robert repeatedly reported little change in his science instruction: "The things I'm doing now I just keep on doing… I like our textbook. I like our interactive workbook". He pointed to the standards, the book, and the materials he had on hand as the primary factors influencing his decisions about science curriculum. In terms of science instruction, he identified time as the main constraint. The structure of the school day left "about an hour a day that is discretionary". The content of instruction during that discretionary time depended on "what kind of work we need to get finished or if we are doing testing and things". For example, in one interview, Robert indicated that with the grading period approaching, "we have to get the writing test done and we have to finish something in math" during the discretionary time that week. He later suggested that "probably the major factor would be when we spend time on social studies instead of science". Robert said that he used a shortened version of SGI about three to five times a year because "it is just time consuming to do it on a regular basis". Each year in interviews, he referred to the same examples of inquiry-based lessons—observing the life cycle of the fruit fly and dissolving salt in water. As noted on Fig. 3, the amount of instructional time Robert devoted to science and his use of inquiry-based and student-centered instructional activities dropped after the program ended and again in the last year. He noted that the school had "a new writing program and that has affected" the time devoted to science. Consequently, science instruction more frequently became a 15-to-20-min textbook-based activity.

Although he "got a lot of good ideas" from other teachers involved in the professional development program, Robert stated that he did not maintain contact with them after the program ended. He initially shared ideas with another second-grade teacher who also participated in the program and who taught in the room next to his. But she retired, leaving him wishing for "at least one other teacher [at his grade level] that wanted to work together". He participated in some school-sponsored professional development on English language development and Common Core strategies in the years after the program ended, but had no additional professional development in science.

As shown in Fig. 3, Robert reported variations in principal support over the years. In interviews, he generally described administrators as supportive of science as long as teachers "get the key things done—and that is reading, writing, and math". The school has a science lab, but Robert did not use it because his science instruction can be done in the classroom, which saves the time of walking students to and from the lab and reserving it in advance. To expand his teaching of science, Robert indicated that he would need "a mandate from the administration that we are supposed to teach science × number of minutes per day… but that's probably not going to happen unless the state says it will be tested" in the early elementary grades.

Summary

Although Robert made gains in self-efficacy in teaching science and appreciated the instructional strategies he learned through the professional development, he changed little about his teaching methods in science. With time constraints, little collegial support, and variations in principal support, he relied heavily on textbook-based instruction after the professional development ended and eventually reduced science to 15–20 min lessons.

Gabriela

Teaching assignment

Throughout the 3-year professional development and 4-year follow-up study, Gabriela taught first grade at a small, rural school that enrolled fewer than 200 students in grades kindergarten to eighth. Given the small school size, there was only one teacher per grade level. The student population consisted of approximately 50% Hispanic and 45% White students. Approximately 70% of the students qualified for free and reduced lunch, and approximately 27% of the students in the school were English language learners.

Change in instructional time

Prior to the professional development, Gabriela reported teaching science about 3 days a week for 10 min or less (see Fig. 4). After the first year, her instructional time increased to lessons of approximately 40–50 min twice a week and then shifted to shorter periods of time on more days of the week. Her reported instructional time in science eventually increased to approximately 50–60 min 5 days a week but then dropped the following year to 20–30 min about three times a week.

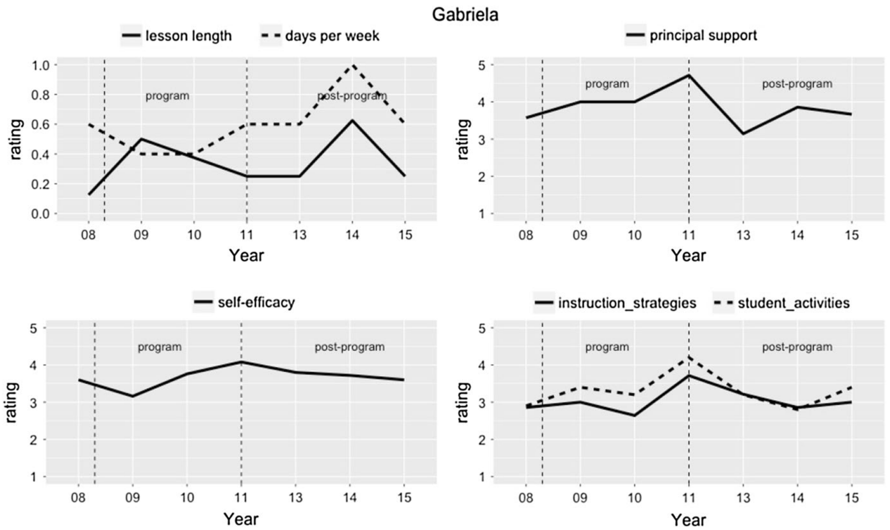

Fig. 4 Change in teacher ratings related to (top left) instructional time: minutes per day in typical lesson (10 or fewer = 0; 11–20 = .13; 21–30 = .25; 31–40 = .38; 41–50 = .5; 51–60 = .625) and days of science instruction per week (none = 0; one = .2; two = .4; three = .6; four = .8; five = 1); (top right) principal support (1 = strongly disagree to 5 = strongly agree); (bottom left) self-efficacy (1 = low to 5 = high); and (bottom right) instructional strategies and student activities (1 = never to 5 = all or almost all)

Impact of professional development

In each interview, Gabriela identified the collaboration and networking with other first grade teachers as a key benefit of the professional development. As the only first grade teacher at her school, she had not been able to work with a grade level partner on science curriculum and instruction. At the professional development meetings, Gabriela loved to talk with other first grade teachers, “pick their brains”, and work together to develop grade level science units. As she described,

> We got more ideas and we got to problem solve and see things differently and say, “Oh, I had never thought of a certain concept in a certain way…”. It was just seeing somebody else’s perspective and seeing how they taught it and what works for them. I really don’t have that [at my school].

When the program ended, she especially missed that collaboration and pointed out: “Now, I’m by myself again”. Although she talked with the second-grade teacher who also participated in the professional development, they did not do joint planning because they “try not to teach the same thing”. Other teachers at her school seemed less invested in teaching science, and “the way they talk about it, they make it seem so hard”. Gabriela “did not want to be dragged down by that”, so she generally opted to “do my own thing’. Despite “negative feedback from other teachers” about the idea, she managed to start a school garden, and she worked with the second-grade teacher to host a school science night for 2-years. When teachers started to resist coming back to school in the evening, the event ended.

Gabriela appreciated other aspects of the professional development such as “the hands-on projects” and “how the scientists break down the information that would normally be very scientific into normal language”. She credited the 3-year program with boosting her confidence in teaching science, particularly physical science, and in using investigative strategies. Her overall self-efficacy and her use of inquiry-based and student-centered instructional strategies increased during the professional development and declined somewhat after the program ended (see Fig. 4).

Factors that affected science instruction

In interviews, Gabriela stated that she continued to use hands-on, inquiry-based lessons developed during the professional development because “they make the science come alive [for the students]… And they love it because it is not just out of a book. It comes alive”. Instead of complaining about the writing involved, her students would ask “When are we going to do science again? When do we get to use our journals?” Beyond student interest, Gabriela viewed the strategies as improving students’ ability to problem solve and willingness to make mistakes: “Once we started doing more science, they got more confidence in working things out and not getting frustrated as easily and being able to work problems out”.

Similar to other teachers, Gabriela identified time as a key constraint to teaching more science: “It seems like we never have enough time to do everything. But I think [limited] time is the worst thing. We actually have to look at the clock, ‘Oh, we have to move on to the next thing’”. To address time constraints, she changed her instructional strategies “only a little bit”. She determined which topics, such as seasons, could be addressed in language arts or social studies and devoted more time in science to topics such as liquids and gasses that were best taught through the SGI model and experiments. With the SGI model, she modified some steps such as limiting the number of objects in the kit inventory, writing out the engaging scenario for the students to read, and providing the first part of prediction sentences. But, for the most part, she opted “to just continue where we left off. It just takes a long, long time to complete a unit”.

Other teachers at her small school did not share her interest in teaching science, but Gabriela benefited from other forms of support at the school level. Although she reported less principal support after the professional development ended (see Fig. 4), Gabriela pointed out that the principal offered some flexibility in the schedule and resource allocation. Each teacher received $250 a year to “use for anything we want”. She chose to use her funding for science. With respect to the schedule, Gabriela stated: “He [the principal] doesn’t tell us you just have to focus on language arts and math. He lets me have the freedom of doing the science”. But the flexibility did not mean that she could devote as much time as she wanted to science instruction: “You have to do your minutes for language arts, for math, and then basically what is left over, that is used for science or social studies or PE (physical education)”. Gabriela observed an increased focus on math and language arts in the school with the implementation of the Common Core standards. In the last year of this research, a change in the language arts schedule led Gabriela to decrease her instructional time in science to 20–30 min about three times a week. Specifically, the school decided to group students and rotate these groups with different teachers for an additional hour of English language development. This change affected her instructional time with students and reduced by half the time she devoted to science.

Because Gabriela sees science as a vital part of the curriculum for her students, she strives to keep teaching it as much as possible: "For me, science is really important, so I have to give them that time". Gabriela stated that she needs to stay optimistic and focus on her students' interest and excitement for science. She summarized: "I just have to remember that what I'm doing is good, and they are learning".

Summary

After the professional development, Gabriela managed to increase her instructional time in science to a high level before it dropped to a moderate amount due to a change in her school's language arts schedule. She wanted to continue using inquiry-based strategies due to the students' interest and the learning that she observed. Though she lacked collegial support in her school, Gabriela benefited from modest funding to use for science, some flexibility in the daily schedule, and her personal drive to teach science.

Cross-case discussion

Our cross-case analysis highlights five key findings related to sustainability of professional development outcomes. First, all four teachers valued the professional development and gained both skills and confidence in teaching science, but all of them also faced constraints. Largely because of students' reactions, the teachers expressed an ongoing desire to teach science using inquiry-based strategies that they learned in the professional development. When using hands-on, inquiry-based strategies, the teachers observed students' increased engagement and excitement about science. Despite their desire to use these instructional approaches, the teachers encountered constraints, particularly stemming from the emphasis on math and language arts. Across schools, the curriculum in the early elementary grades focused primarily on math and language arts, and the standardized testing focused exclusively on those two areas. In one case, the school implemented a new writing program that further decreased instructional time for science. At another school, the principal, due to the school's status as a program improvement site, contracted with a program to try and raise student test scores in math and language arts. In some schools, teachers had more discretionary instructional time that could be devoted to science. Beyond the emphasis on math and language arts, teachers encountered varied levels of resources, collaboration with colleagues, parent volunteers, and administrative support for teaching science

Second, constraints on instructional time led to less frequent or adapted use of inquiry-based strategies over time. Despite teachers' positive views about SGI, they struggled to find instructional time to implement the full model, which included nine steps. Two of the teachers, Rachel and Gabriela, modified or eliminated some of the steps and trimmed the amount of writing for the early elementary students in order to reduce the overall amount of time needed. Miranda reported using all of the steps but extending the time frame to complete the model, sometimes from days to weeks. Robert opted to use a shortened version of SGI for specific science topics several times a year, but he eventually reverted to frequent use of short textbook lessons in science. Only one of the four teachers ended up dropping inquiry-based strategies, but all of them reported changes in their use of these strategies over time.

Third, factors that influenced the instructional day affected science instruction in both positive and negative ways. For instance, strict pacing schedules at some schools left little time for science instruction, particularly on a consistent basis. As described above, teachers had less curricular flexibility at schools with increased pressure to raise student test scores. In contrast, at schools without strict daily schedules, teachers could more easily fit science into their instructional plans, sometimes by integrating science with other subjects. The most unusual factor that facilitated science instruction for one teacher was the bus schedule. Since there was only one school bus run at this small, kindergarten through eighth grade rural school, all grades had the same instructional day. Consequently, Rachel had an extra hour a day, which enabled her to include more science instruction. These variations in the instructional day reflect the varying amounts of instructional time in science that teachers reported over the years.

Fourth, the professional development program bolstered individual level factors such as teachers' skills and self-efficacy in science, but school level factors, such as principal support and resources, shifted over time and often worked against sustaining instructional changes. The plots for the four teachers provide a visual representation of the shifts in their reported support from school principals. Over the 7-years of this study, all of the teachers worked with multiple principals. As new principals arrived, they often brought differing priorities and allocations of resources, which had both positive and negative effects on teachers' science instruction. For example, a new principal at Miranda's school ended up switching the family science night to a family literacy night. Gabriela benefited from an allocation of $250 that she opted to use on science, but she ended up reducing science instruction because of a change in the language arts schedule. Perhaps more influential on teachers' science instruction was the amount of flexibility that principals afforded teachers in terms of curriculum and the instructional day. Rachel encountered differing levels of support as principals changed over the years, but, with each one, she retained the ability to determine her daily instructional schedule and the option to integrate science with other subjects. This flexibility, combined with other school level factors, led her to spend more instructional time on science than the other case study teachers.

Fifth, teachers with high personal commitment to teach science found ways to cope with constraints, but personal interest and coping mechanisms may not suffice over the long term. The teachers who described personal motivation and interest were more likely to find ways to teach science and to continue using inquiry-based methods in their instruction. For example, Miranda highlighted her own fascination with science and described wanting to help students experience a similar passion for the subject. But she encountered contextual challenges such as the school's designation as a program improvement site, a strict pacing schedule, a new program focused on raising students' test scores in math and language arts, and four different principals in 7-years. Despite these obstacles, she managed to continue using investigative strategies after the program ended and to slightly increase instructional time in science. In contrast, Robert referred to the textbook and standards as key influences on his decisions about teaching science. He noted students' interest in hands-on, inquiry-based activities, but over time, he ended up focusing more on math and language arts instruction and relying on short, 15-min, textbook-based activities in science. Without contextual supports or personal drive, Robert found it too challenging to sustain science instruction over time. Without contextual supports but with personal drive, Miranda continued to teach science in her second-grade classroom. But she modified her instructional strategies and taught science less frequently than other teachers.

Conclusion and implications

This research examined reported changes in early elementary teachers' science instruction over a 7-year period that included their participation in a 3-year professional development program. The case studies of four teachers who had differing patterns of instructional time offer insights into the reasons why instructional shifts occur after professional development ends. Our findings highlight three key issues related to sustaining professional development outcomes in science.

First, these case studies suggest that contextual factors are as important as individual factors in sustaining science instruction after professional development ends. A central focus of professional development programs is building teachers' skills and individual abilities in order to improve classroom instruction in science (Banilower et al. 2007; Heck and Crawford 2004; Heck et al. 2006). Given the typically limited STEM backgrounds of rural teachers (Avery 2013; Arnold et al. 2004) and the challenge in providing content-specific professional development in rural schools (Harmon et al. 2007), changes in teachers' abilities and confidence in teaching science are critical. However, as this study illustrates, professional development may succeed in building teachers' skills and self-efficacy in science teaching, but without supportive contextual factors, science instruction in the elementary grades is unlikely to be sustained over time. These case studies highlight how contextual factors both promoted and inhibited instructional time in science as well as teachers' use of inquiry-based instructional strategies.

Second, the most challenging contextual constraints for teachers were a lack of instructional time and administrative support, stemming from competing curricular priorities, and limited ongoing opportunities to collaborate with other teachers. For the early elementary grades, the emphasis on math and language arts was particularly influential across schools. But, at some schools, factors such as state designation as a program improvement site brought new programs and increased emphasis on those subjects. School principals also changed over the years; and, reflecting the difficult decisions facing administrators in rural settings (Farmer 2009), the changes also brought shifts in administrative priorities and support for science education. In addition, given the small school sizes, the teachers often had few or no grade-level colleagues at their school to collaborate with on developing inquiry-based science instruction. The professional development offered regular opportunities for participants to collaborate with grade-level colleagues, but, given the distances between schools, those connections became more challenging to maintain after the professional development ended. Although the teachers maintained their ability and interest in using investigative and inquiry-based strategies to teach science, they reported less frequent and modified use over time due to contextual challenges that were heightened in their rural settings. These contextual supports and constraints fluctuated not only across schools but also from year to year.

Third, relying on teachers' personal motivation to teach science in the face of ongoing contextual challenges is not a viable approach for sustaining science instruction over time. The case studies illustrate how an individual teacher's interest in science can be a positive contributing factor in decisions about instructional time and strategies. But personal motivation cannot compensate fully for a lack of supports at the school level. In rural schools, the competing and shifting contextual challenges that teachers encounter (Jimerson 2005) combined with limited organizational support (Harmon et al. 2007), stand to overwhelm the interest in science engendered by professional development. Rural elementary schools that are able to maintain high instructional time in science after professional development ends have supportive school contexts combined with teachers motivated by student interest in science (Sandholtz and Ringstaff 2016). When lacking supports, some teachers will find a way to include science in their curriculum, but they simply cannot allocate the same amount of instructional time as teachers in more supportive settings, whether in rural communities or other settings. Contextual constraints that continue over the long-term may gradually erode teachers' commitment to teaching science, especially in the early elementary grades where science is not a required part of the curriculum and in rural settings where teachers are more isolated and have less access to science-focused educational opportunities (Avery 2013).

Professional development plays an important role in science education reform efforts. But professional development outcomes are unlikely to persist over time without contextual supports for science instruction. Teachers' professional skill in teaching science may be thwarted by policies and accountability measures that emphasize mathematics and language arts and fail to include science in the early elementary grades (Hayes and Trexler 2016; Judson 2013). Without policies that make science as important as mathematics and language arts in the early elementary grades, changes in science instruction stemming from professional development will continue to be challenging to sustain. Although those providing professional development are typically not in a position to make changes to policies and other contextual factors, they may be able to offer some supports for teachers such as follow-up sessions, materials and supplies, or other science-focused opportunities. More longitudinal studies that examine the influence of context on elementary teachers' decisions about instructional time and strategies in science are important to determine the types of supports that are workable and will promote sustainability of professional development outcomes.

Acknowledgements This material is based on work supported by the National Science Foundation under Grant No. DRL-1119589. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Appleton, K., & Kindt, I. (1999). Why teach primary science? Influences on beginning teachers' practices. *International Journal of Science Education, 21*(2), 155–168.

Arnold, M., Gaddy, B., & Dean, C. (2004). *A look at the condition of rural education research: Setting a direction for future research*. Aurora, CO: Mid-continent Research for Education and Learning.

Avery, L. M. (2013). Rural science education: Valuing local knowledge. *Theory into Practice, 52*(1), 28–35.

Banilower, E. R., Heck, D. J., & Weiss, I. R. (2007). Can professional development make the vision of the standards a reality? The impact of the National Science Foundation's Local Systemic Change through Teacher Enhancement Initiative. *Journal of Research in Science Teaching, 44*(3), 375–395.

Bogdan, R. C., & Biklen, S. K. (1998). *Qualitative research for education: An introduction to theory and methods* (3rd ed.). Boston: Allyn & Bacon.

Bowes, A. S., & Banilower, E. R. (2004). *LSC observational study: An analysis of data collected between 1998 and 2003*. Chapel Hill, NC: Horizon Research Inc.

Creswell, J. W. (2005). *Educational research. Planning, conducting, and evaluating quantitative and qualitative research* (2nd ed.). Upper Saddle River, NJ: Pearson.

Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, N., & Orphanos, S. (2009). State of the profession: Study measures status of professional development. *Journal of Staff Development, 30*(2), 42–50.

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

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academies Press.

Farmer, T. A. (2009). Unique rural district politics. *The Rural Educator, 30*(2), 29–33.

Flyvberg, B. (2001). *Making social science matter*. Cambridge: Cambridge University Press.

Franke, M. L., Carpenter, T. P., Levi, L., & Fennema, E. (2001). Capturing teachers' generative change: A follow-up study of professional development in mathematics. *American Educational Research Journal, 38*(3), 653–689.

Gaikhorst, L., Beishuizen, J., Zijlstra, B., & Volman, M. (2017). The sustainability of a teacher professional development programme for beginning urban teachers. *Cambridge journal of education, 47*(1), 135–154.

Gess-Newsome, J. (2001). The professional development of science teachers for science education reform: A review of the research. In J. Rhoton & P. Bowers (Eds.), *Professional development: Planning and design* (pp. 91–100). Arlington, VA: NSTA Press.

Guskey, T. R. (2002). Does it make a difference? Evaluating professional development. *Educational Leadership, 59*(6), 45–51.

Guskey, T. R., & Sparks, D. (2002). *Linking professional development to improvements in student learning*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Haney, J. J., Lumpe, A. T., Czeriak, C. M., & Egan, V. (2002). From beliefs to actions: The beliefs and actions of teachers implementing change. *Journal of Science Teacher Education, 13*(3), 171–187.

Harmon, H., Gordanier, J., Henry, L., & George, A. (2007). Changing teaching practices in rural schools. *The Rural Educator, 28*(2), 8–12.

Harmon, H., & Smith, K. (2007). *A legacy of leadership and lessons learned: Results from the rural systemic initiatives for improving mathematics and science education*. Charleston, WV: Edvantia Inc.

Hawley, W. D., & Valli, L. (1999). The essentials of effective professional development: A new consensus. In L. Darling Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 127–150). San Francisco: Jossey-Bass.

Hayes, K. N., & Trexler, C. J. (2016). Testing predictors of instructional practice in elementary science education: The significant role of accountability. *Science Education, 100*(2), 266–289.

Heck, D. J., & Crawford, R. A. (2004). *LSC teacher questionnaire study: A longitudinal analysis of data collected between 1997 and 2003*. Chapel Hill, NC: Horizon Research Inc.

Heck, D. J., Rosenberg, S. L., & Crawford, R. A. (2006). *LSC teacher questionnaire study: A longitudinal analysis of data collected between 1997 and 2006*. Chapel Hill, NC: Horizon Research Inc.

Horizon Research, Inc. (2000). *Local systemic change through teacher enhancement science teacher questionnaire.* Chapel Hill, NC: Horizon Research, Inc. Retrieved from http://www.horizon-research.com/instruments/lsc/tq_k8sci.php.

Jimerson, L. (2005). Special challenges of the "No Child Left Behind" Act for rural schools and district. *The Rural Educator, 26*(3), 1–4.

Johnson, C. (2006). Effective professional development and change in practice: Barriers science teachers encounter and implications for reform. *School Science and Mathematics, 106*(3), 150–161.

Judson, E. (2013). The relationship between time allocated for science in elementary schools and state accountability policies. *Science Education, 97*(4), 621–636.

McLaughlin, M., & Mitra, D. (2001). Theory-based change and change-based theory: Going deeper, going broader. *Journal of Educational Change, 2*(4), 301–323.

Monk, D. H. (2007). Recruiting and retaining high-quality teachers in rural areas. *The Future of Children, 17*(1), 155–174.

National Academy of Sciences. (2015). *Science teachers' learning: Enhancing opportunities, creating supportive contexts*. Washington, DC: The National Academies Press.

National Staff Development Council. (2001). *Standards for staff development*. Oxford, OH: National Staff Development Council (NSDC).

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Pea, C. H. (2012). Inquiry-based instruction: Does school environmental context matter? *Science Educator, 21*(1), 37–43.

Penuel, W., Fishman, B., Gallagher, L., Korbak, C., & Lopez-Prado, B. (2008). Is alignment enough: Investigating the effects of state policies and professional development on science curriculum implementation. *Science Education, 93,* 656–677.

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

Penuel, W. R., & Gallagher, L. P. (2009). Preparing teachers to design instruction for deep understanding in middle school Earth science. *Journal of the Learning Sciences, 18*(4), 461–508.

Riggs, I., & Enochs, L. (1990). Development of an elementary teacher's science teaching efficacy belief instrument. *Science Education, 74,* 625–637.

Rotermund, S., DeRoche, J., & Ottem, R. (2017). *Teacher professional development by selected teacher and school characteristics: 2011–12*. Washington, DC: U.S. Department of Education.

Sandholtz, J. H., & Ringstaff, C. (2013a). Assessing the impact of teacher professional development on science instruction in the early elementary grades in rural US schools. *Professional Development in Education*, *39*(5), 678–697.

Sandholtz, J. H., & Ringstaff, C. (2013b). *Does professional development make a difference? Results of a 3-year study of K-2 science instruction and student learning*. Paper presented at annual meeting of the American Educational Research Association, San Francisco, CA.

Sandholtz, J. H., & Ringstaff, C. (2016). The influence of contextual factors on the sustainability of professional development outcomes. *Journal of Science Teacher Education*, *27*(2), 205–226.

Settlage, J., & Meadows, L. (2002). Standards-based reform and its unintended consequences: Implications for science education within America's urban schools. *Journal of Research in Science Teaching, 39*(2), 114–127.

Settlage, J., Southerland, S., Smetana, L., & Lottero-Purdue, P. (2017). *Teaching science to every child: Using culture as a starting point* (3rd ed.). New York: Routledge.

Sherry, L. (2002). Sustainability of innovations. *Journal of Interactive Learning Research, 13*(3), 211–238.

Sparks, D. (2002). *Designing powerful professional development for teachers and principals*. Oxford, OH: National Staff Development Council.

Spillane, J. P., & Callahan, K. A. (2000). Implementing state standards for science education: What district policy makers make of the Hoopla. *Journal of Research in Science Teaching, 37*(5), 401–425.

Steele, D. F. (2001). The interfacing of preservice and inservice experiences of reform-based teaching: A longitudinal study. *Journal of Mathematics Teacher Education, 4,* 139–172.

Stigler, J., & Hiebert, J. (1999). *The teaching gap*. New York: Free Press.

Sullivan-Watts, B. K., Nowicki, B. L., Shim, M. K., & Young, B. J. (2013). Sustaining reform-based science teaching of preservice and inservice elementary school teachers. *Journal of Science Teacher Education, 24*(5), 879–905.

Supovitz, J. A., & Turner, H. M. (2000). The effects of professional development on science teaching and practices and classroom culture. *Journal of Research in Science Teaching, 37*(9), 963–980.

Vanosdall, R., Klentschy, M., Hedges, L. V., & Weisbaum, K. S. (2007). *A randomized study of the effects of scaffolded guided-inquiry instruction on student achievement in science*. Paper presented at the conference of the American Educational Research Association, Chicago, IL.

Wyner, Y. (2013). The impact of a novel curriculum on secondary biology teachers' dispositions toward using authentic data and media in their human impact and ecology lessons. *Journal of Science Teacher Education, 24*(5), 833–857.

Yin, R. K. (2003). *Case study research design and methods* (3rd ed.). Thousand Oaks, CA: Sage.

Zehetmeier, S., & Krainer, K. (2011). Ways of promoting the sustainability of mathematics teachers' professional development. *ZDM Mathematics Education, 43,* 875–887.