Routledge
Taylor & Francis Group

Journal of Science Teacher Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/uste20

Evaluating the Effectiveness of an Online Professional Development for Rural Middle-School Science Teachers

Brooke A. Moore, Earl F. Legleiter, Kylia Owens, Brynne Packard & Jessica Wright

To cite this article: Brooke A. Moore, Earl F. Legleiter, Kylia Owens, Brynne Packard & Jessica Wright (28 Nov 2023): Evaluating the Effectiveness of an Online Professional Development for Rural Middle-School Science Teachers, Journal of Science Teacher Education, DOI: 10.1080/1046560X.2023.2287321

To link to this article: https://doi.org/10.1080/1046560X.2023.2287321

Published online: 28 Nov 2023.

JOURNAL OF SCIENCE TEACHER EDUCATION
https://doi.org/10.1080/1046560X.2023.2287321

Evaluating the Effectiveness of an Online Professional Development for Rural Middle-School Science Teachers

Brooke A. Moore, Earl F. Legleiter, Kylia Owens, Brynne Packard, and Jessica Wright

College of Education, Fort Hays State University, Hays, Kansas, US

ABSTRACT

In rural, geographically dispersed school districts, access to high-quality face-to-face professional development (PD) is challenging. Our study developed and compared the effectiveness of an online PD for middle-school science teachers working in remote, rural areas of Kansas with an evidence-based traditional face-to-face PD with the goal of supporting change in teachers' conceptual understanding and self-efficacy in utilizing Next Generation Science Standards (NGSS), change in instructional practices, and overall student content learning. Fifteen rural middle-school science teacher participants were randomly selected to participate in the online or traditional PD, then taught the *Toward High School Biology* curriculum to their 504 middle-school students. Findings aligned with our hypothesis that online PD is as effective as traditional in improving student content knowledge related to the curriculum. Teachers' instructional practices in using Next Generation Science Standards improved, as did their use of student-centered instruction and making science relevant to the lives of their students. Our study supports the importance of designing PD that effectively and economically brings teachers in isolated and diverse rural communities together.

KEYWORDS

Middle-school; professional development; rural education; science education

Introduction

High-quality professional development (PD) is a proven approach in supporting teachers. Content-focused PD, specifically in the science, technology, engineering, and math (STEM) field, increases the academic achievement of students (Barrett et al., 2015; Hjalmarson et al., 2020; Margot & Kettler, 2019). Most PD opportunities provided to STEM educators are offered through traditional face-to-face methods. In rural areas, accessing high-quality, relevant PD can pose a challenge. Physical distance from universities and other outside providers makes access to PD a significant barrier. Recognizing the value and benefit of PD on fostering teachers' use of evidence-based instructional approaches in science education, we argue that it is important to design PD that effectively and economically brings teachers in isolated and diverse rural communities together to improve science instruction. Our National Science Foundation (NSF)-funded research study specifically targeted middle-school science teachers working in rural communities in Kansas and serving students with high needs (i.e., living in poverty, English learners, culturally diverse, migrant). We developed an online PD model and then assessed its effectiveness as compared to an aligned face-

CONTACT Brooke A. Moore brooke.moore@fhsu.edu College of Education, Fort Hays State University, 600 Park Street, Hays, KS 67602

to-face (F2F) PD in supporting change in teacher conceptual understanding, self-efficacy in utilizing *Next Generation Science Standards* (NGSS; NGSS Lead States, 2013), change in instructional practices, and student content learning, following participation in our matched PDs and students' instruction in an NGSS-aligned curriculum.

Review of the literature

The National Center for Education Statistics (NCES) locale code system denotes three rural categories: fringe (less than or equal to five miles from an urban area), distant (between 5 and 25 miles from an urban area), and remote (beyond 25 miles from an urban area; NCES, 2019). The US Department of Agriculture (USDA) visually defines rural as "open countryside with population densities of less than 500 people per square mile and places with fewer than 2,500 people" (USDA, 2019). As Showalter et al. (2019) has argued, individuals living in rural places in the US are not getting the attention that is needed to address the challenges they face. This is vital as there are more than 9.3 million US children (e.g., one in five) attending rural schools. Of those 9.3 million rural children, one in six is living below the poverty line, one in seven qualify for special education services, and one in nine are from migrant families (Showalter et al., 2019).

In their work in rural settings, Barrett et al. (2015) found that rural school districts rely heavily on local teachers who are trained in nearby universities, and once hired by the district, tend to remain in that district for their careers. Their argument is that rural school districts often lack the resources to recruit much from beyond their setting, so providing high-quality PD to their faculty is vital to ensure equitable student achievement (National Research Council, 2015). Yet, rural districts face significant challenges in providing teachers with high-quality PD. Along with being widely dispersed geographically across parts of the US, the programming developed for metropolitan educators may not be relevant to the needs of educators working in rural settings (Lavalley, 2018). Williams (2012) noted that isolated rural locations have a high concentration of children living in poverty and African American, Native American, or nonwhite Hispanic students. In Kansas, where this study took place, half of all schools are in rural communities (Showalter et al., 2019). Kansas Science Assessment results illustrate the disparity that exists between rural and other schools. Science scores for Kansas schools, the target area for this research study, showed only 23.5% of students scoring at level 4 (performing above expectations) compared to 43.5% state-wide. In addition, southwest Kansas has high numbers of English learners, students from migrant families, and a higher-than-average number of students who qualify for free or reduced lunch (63.79% versus the state average of 48.99%). This confluence of low economic, limited English proficiency, and migrant status of students introduces a complex challenge for teachers working to fill in the gaps in student background knowledge while simultaneously increasing content understanding.

Research consistently demonstrates the positive impact of high-quality teacher professional development to train new teachers and support continuing education for experienced teachers, especially in STEM fields. In fact, math- and science-focused PD has been shown to increase the academic achievement of students in the year that they are assigned to the target teacher, and throughout the following school year (Barrett et al., 2015). In their seminal work, Darling-Hammond et al. (2017) identified key features of effective PD: content focused, involves active learning, supports collaboration, uses models of effective instructional practice, provides coaching and expert support, offers feedback and reflection, and is sustained over time. Althauser (2015) noted that intensive, sustained, job-embedded PD is more likely to improve teacher's content knowledge, instructional practices, and student achievement. Job-embedded PD provides learning within the context of a - teacher's day-to-day instructional practice.

Any discussion of how teachers will increase student achievement in science must be anchored in the new vision of science teaching and learning. A Framework for K-12 Science Education (NRC, 2015) and the Next Generation Science Standards (NGSS; NGSS Lead States, 2013) describe objectives for students' learning based on key insights from research on STEM education. Currently, many Kansas teachers have not been taught using the new three-dimensional learning approaches for implementing the NGSS (e.g., science and engineering practices, crosscutting themes, disciplinary core ideas); therefore, significant PD is needed (NRC, 2015). To implement NGSS, teachers must reconsider the science content that is taught, how students build their understanding of that content, and how ideas fit together to tell a coherent story (Reiser, 2013).

To date, more is being learned about the role of online learning for practicing educators. Lantz-Anderson et al. (2011) reviewed 20 years of literature on online teacher communities and found benefit from the opportunity of teachers to connect more broadly with other professionals through sharing of knowledge and experiences. Noting that access to high-quality PD is vital but limited for some, Powell and Bodur (2010) conducted a survey of six high school educators after engaging in an online PD. Findings indicated key features of successful online PD: relevancy and authenticity to their own contexts, usefulness of the content, embedded interaction and collaboration with other participants, and time for reflection on their own learning. These findings match the work of Blitz (2013) who conducted a literature review on the effectiveness of online professional learning communities, noting the importance of embedding structures for collaboration, allowing participants to shape their learning goals, pairing participants with experts, and designing self-reflection activities. Fishman et al. (2013) compared teacher and student learning across an online and traditional F2F PD, noting that both conditions improved in content knowledge. There were no differences between PD settings on teacher self-efficacy or beliefs. The authors concluded that if there are no perceived differences between traditional PD and online, the cost of developing online makes sense, particularly for teachers in more remote areas.

While the current literature indicates that online PD is as effective as traditional PD, recent research , specifically, in rural locales is somewhat sparse. Erickson et al. (2012) examined the effectiveness of an asynchronous online PD for special educators in rural and non-rural settings, with no comparison of a traditional PD. Findings indicated similar gains for both groups of participants in content knowledge, use of evidence-based practices, and implementation of special education services. Hunt-Barron et al. (2015) used digital tools (e.g., blogging, Google Sites) to enhance a traditional face-to-face PD for Language Arts teachers to offer continued support after the PD ended. Findings indicated that some participants engaged by reporting on implementation strategies and reflecting on positive outcomes learned in the PD. Finally, Moser and Wei (2023) examined the perceptions of 50 English as a Second or Other Language (ESOL) teachers following an online PD in improving their knowledge and practices in working with English learners in rural contexts.

These studies support the evidence that online PD is effective in rural settings yet lack randomized control designs to compare with a traditional F2F (Fishman et al., 2013). Furthermore, minimal evidence indicates the impact for science educators. Our goal was to develop an online version of a proven F2F PD to better meet the needs of rural district science teacher professional development.

More recent research in the development of effective online PD notes the importance of designing online learning differently from more traditional face-to-face settings, a call made by Blitz (2013), yet retaining many of the key features that prove relevant for effecting teacher change. Bragg et al. (2021) conducted a review of the literature to identify effective online PD practices, specifically noting the importance of a focus on design elements that support participants' engagement and learning, allow for individuality in learning styles, focus on the acquisition of content knowledge, encourage application of new learning, and foster collaborative engagement. Design elements that promote engagement and learning include the use of gamification design (Deterding et al., 2011) such as levels, turn taking, goals, and play-centric features (e.g., earning points, conquering obstacles). Tseng and Kuo (2014) examined social participation as a feature of effective online learning through a survey of 321 teachers engaged in an online professional learning community. The authors found that when members felt a closer connection to one another, they were more likely to share knowledge and help others. From their work, the authors suggested that online PD focus on building interpersonal relationships and increasing the efficacy of users' willingness to share and participate more fully. Wasserman and Migdal (2019) compared perceptions of teachers enrolled in online versus traditional PD courses. When online PD included key features of successful online learning, participants indicated that they felt connected to the context and content despite a lack of physical presence. Therefore, the goal of our project was to design an effective online PD that matched a proven, traditional face-to-face PD to determine the effectiveness for middle-school science teachers working in rural settings. Noting the importance of high-quality science instruction that aligns with the NGSS science standards, our goal was to find ways to deliver PD that is impactful to teachers (e.g., boosts content knowledge, self-efficacy, and use of NGSS aligned instructional practices) and cost-effective to their school districts.

Conceptual framework and logic model

For our study, we applied the interactive, constructive, active, and passive (ICAP) framework (Chi & Wylie, 2014). The ICAP framework identifies four levels of overt learning behaviors linked to ever-increasing stages of conceptual understanding (Chi & Wylie, 2014). *Passive* engagement is characterized by participants receiving information (i.e., listening to a lector), leading to simple recall of what was learned. *Active* engagement involves physical manipulation of the learning material (i.e., taking notes, examining an object) that is integrated into a knowledge schema and can then be applied to similar situations. *Constructive* engagement occurs when the individual generates something new (i.e., draws a diagram, builds a model) that can relate to other ideas and be transferred to new situations. *Interactive* engagement is the cognitive level where individuals are engaged in generative dialogue with others to co-create new knowledge. By using ICAP, we sought to generate deep conceptual knowledge related to science and the use of NGSS regardless of the PD setting (i.e., F2F, online).

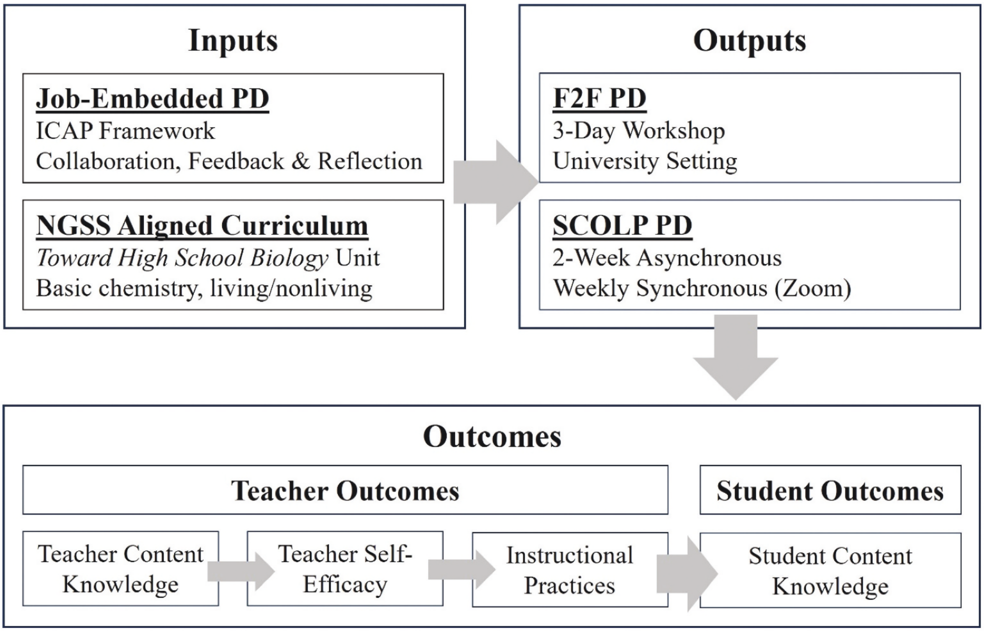

Figure 1. Logic model.

Our logic model (see Figure 1) begins with two primary inputs: job-embedded PD for teachers and NGSS-aligned curriculum. Both our F2F and online PD experiences used the ICAP framework, including elements that were passive, active, constructive, and interactive. We also drew from the research on effective PD to ensure both settings included collaboration, a chance to try out instructional practices with feedback, and opportunities for reflection specifically related to utilizing NGSS three-dimensions of learning with their own students (Althauser, 2015; Darling-Hammond et al., 2017). As noted by Althauser (2015), job-embedded PD must be focused on content instruction (e.g., curriculum specific), instructional methods (e.g., NGSS three-dimensional learning), and assessment (e.g., students pre- and post-unit content knowledge assessments). To ensure both PDs were content-focused and related to participants' everyday practice of teaching middle-school science, we utilized the *Toward High School Biology* (*THSB*) curriculum unit, aligned to NGSS (Herrmann-Abell et al., 2016). The *THSB* curriculum was developed to use NGSS three-dimensions (e.g., science and engineering practices, crosscutting concepts, and disciplinary core ideas) to guide middle-school students in applying basic chemistry ideas to living and nonliving contexts. In examining *THSB's* effectiveness, Herrmann-Abell et al. (2016) found that students who were taught using the *THSB* unit performed better on posttest outcome measures and demonstrated fewer misconceptions related to the content than peers who were in business-as-usual science classrooms. In both PD settings, modeling of NGSS three-dimensions of learning occurred as teachers experienced the unit as if they were students by reading science texts, building molecular models, analyzing data, and constructing scientific explanations by applying disciplinary core ideas. Each PD was designed with an introduction to the NGSS science standards followed by sections aligned to the four chapters in the *THSB* curriculum. Regardless of PD setting, as participants engaged in the chapters, they moved through the curriculum as their students would—taking consideration of how they would teach the lessons in their own contexts, what supports they would provide for diverse and exceptional learners, and learning how to analyze student work to better understand student misconceptions, thus ensuring the PD was job-embedded. The content of both PD settings came directly from the *THSB* curriculum (e.g., lessons, videos).

Our logic model outputs include our matched PD settings (F2F, online). We utilized the traditional F2F model PD that was developed by the creators of the *THSB* curriculum as our foundation. Working with the curriculum development team which included the *THSB* authors, we first mapped out our F2F PD to include three full-day sessions during the summer, hosted at a university in western Kansas. During the PD, participating teachers were introduced to the *THSB* and NGSS aligned practices by first experiencing a lesson from the unit as students. Then, participants learned about NGSS three-dimensions, as embedded within the curriculum. Using our ICAP framework (Chi & Wylie, 2014), over the 3 days, participants worked in small groups, listening to presenters' lectures (passive participation), examining student expectations within the unit (active participation), building molecular models aligned with the unit using LEGO and ball and stick modeling kits (constructive participation), and working in their small groups to plan lessons and reflect on their own learning (interactive participation). See Table 1 for more details on how the F2F and SCOLP PD contexts were aligned.

Our second output was developing our online PD which we called the Science Comprehensive Online Learning Platform (SCOLP). As Blitz (2013) argued, effective online cannot just convert a traditional PD into an online PD as there are problems inherent in

Table 1. Side-by-side comparison of the PD conditions.

PD Features	F2F	SCOLP
PD Overview	Three-day workshop format with facilitators presenting information to the whole group, including an overview of NGSS three-dimensional learning and the four chapters in the *THSB* curriculum.	Three-week workshop where participants experience levels which align to the PD overview (e.g., what is NGSS, why this is important) and the four chapters in the THSB curriculum.
Asynchronous	Participants sit in small groups and are assigned quiet time in the PD to read lesson assignments in the THSB curriculum, record observations on their own as they watch videos of experiments, build molecular models using ball & stick or LEGO modeling kits, review student work, and develop lesson plans. Participants complete homework assignments each night.	*Most of the content of the PD is presented asynchronously so participants can work through at their own pace.* Participants work through levels completing tasks to earn Experience Points (XP). Tasks match the asynchronous activities outlined in the F2F PD context (e.g., video observations, modeling, etc.).
Synchronous	*Most of the PD content is presented synchronously by the facilitators through lecture and video and small group work.* Facilitators monitor individual work and support small groups in discussing observations, student work, and developed lessons. Facilitators review homework assignments and discuss in whole group ways to implement *THSB*, including supports for diverse and exceptional learners.	Facilitators review participant responses to the interactive slides, discussion posts, and homework, then meet at the end of each week in the PD to discuss questions that have emerged during asynchronous work and engage in planning for implementation of the THSB, including supports for diverse and exceptional learners.
Assessing Teacher Learning	Homework assignments Informal Small Group Discussions Teacher Developed Lesson Plans Pre- and Post-PD Assessments	Interactive Participation Discussion Responses to Prompts/ Responses to Colleagues in PD Experience Points (XP) Earned Teacher Developed Lesson Plans Pre- and Post PD Assessments

mapping real-time, F2F learning onto asynchronous settings. More attention must be dedicated to the features of online learning that are most effective. From the research examining online PD, we recognized that an effective and engaging online PD must provide learners with experiences that transcend content delivery and summative assessment. Designing collaborative online PD shifts emphasis from the use of individual technology to individuals using technology to support relationships (Siemens, 2005). Online PD should provide community involvement and networking opportunities where learners are able to actively participate as they contribute to the process as authors and knowledge managers (Ivanova & Popova, 2011; Tseng & Kuo, 2014). Our online platform, SCOLP, was designed using Stream LXP technology (formally known as *Curatr*). This platform provided rich opportunities for our participants to engage asynchronously with the content and synchronously with each other in discussions and sharing of ideas. Participants who were randomly assigned to SCOLP completed the online PD over the course of 3 weeks during the summer. They worked through units of the *THSB* curriculum asynchronously, watching videos and reading content (passive participation), examining student expectations within the unit (active participation), and building molecular models aligned with the unit using LEGO and ball and stick modeling kits (constructive participation). Participants then met in small groups synchronously through Zoom each week with a project team member to support understanding, application of the unit to their contexts, and reflect on their own learning and instructional practices (interactive participation). Figure 2 provides screen shots of SCOLP online through Stream LXP.

Our logic model outcomes are based on our hypothesis that, following effective F2F or online PD, we would see improvements in teacher content knowledge which would lead to enhanced teacher self-efficacy which would lead to change in instructional practices which would ultimately support overall growth in students' content knowledge.

Methods

Research questions

Based on the existing literature (Blitz, 2013; Fishman et al., 2013; Wasserman & Migdal, 2019), we hypothesized that our online PD (SCOLP) would be as effective as our traditional F2F PD in fostering the sustained implementation of evidence-based science reform instructional practices, NGSS, that would lead to gains in student learning. We hypothesized our online PD, SCOLP, to be an effective approach to supporting deep conceptual understanding of the *THSB* unit. We also hypothesized that an overall improvement in teacher content knowledge, teacher self-efficacy, and change in instructional practice would lead to significant growth in student content knowledge of the *THSB* unit. For our study, we examined differences within and across teacher groups with our first research question and within and across student groups with our second research question:

(1) What differences exist within and between the conditions (F2F, SCOLP) in teacher content knowledge, teacher self-efficacy, and change in instructional practices following the PD and teaching the *THSB* unit 2 years in a row?
(2) What differences exist within and between the conditions (F2F, SCOLP) in student content knowledge following instruction in the *THSB* unit?

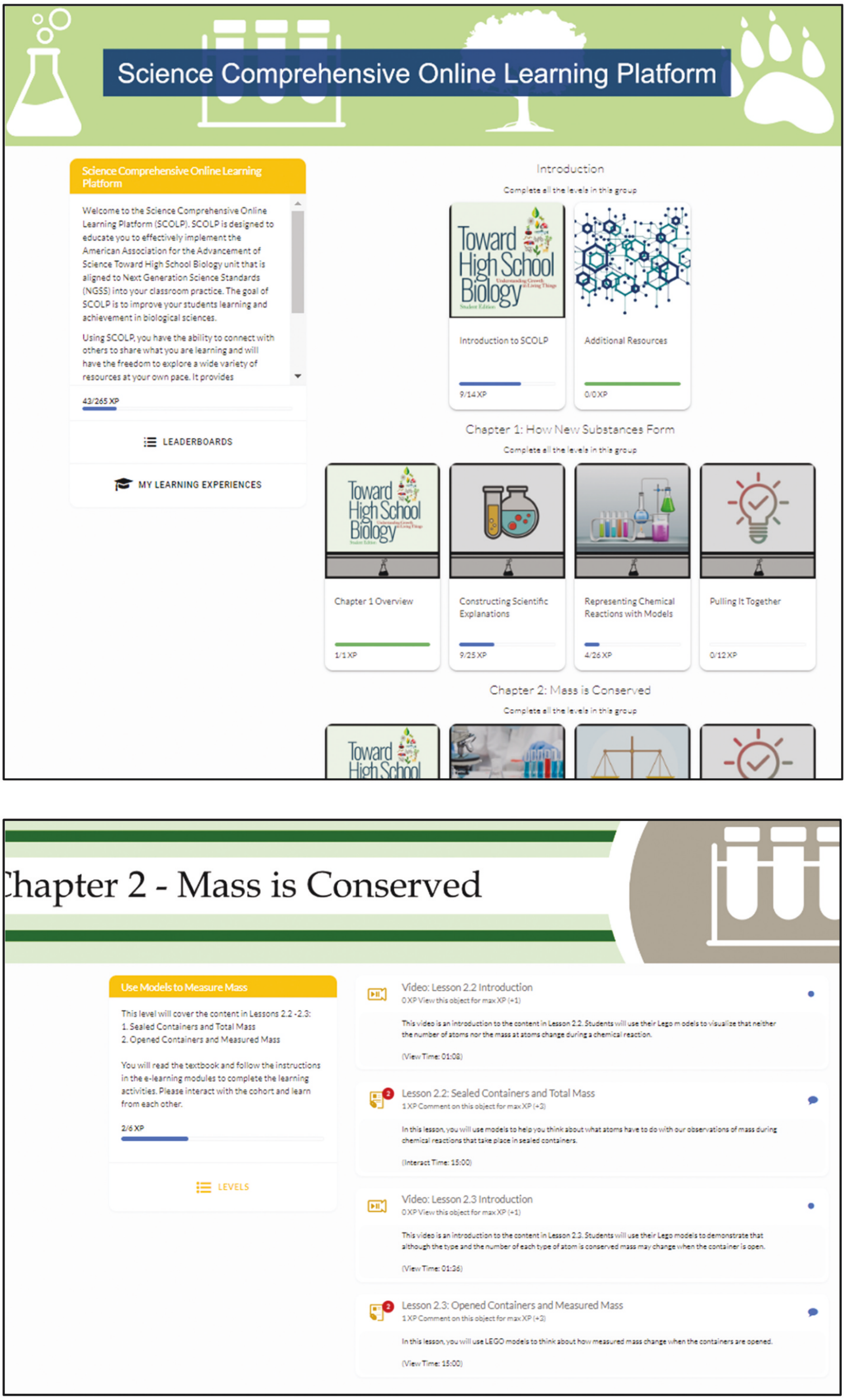

Figure 2. Screen shots of the SCOLP PD context.

Setting and participants

Knowing that rural districts face challenges in providing teachers with high-quality PD, we purposefully recruited middle-school science teachers teaching either seventh or eighth grade from rural districts across portions of southwest Kansas. As our participating rural districts were small, the recruited science teacher was often the only middle- and high-school science teacher in the district. Teachers were then randomly assigned to either our F2F PD or our SCOLP PD.

While we were initially able to recruit and consent 41 middle-school science teachers, only 27 completed their assigned PD and taught the *THSB* unit in year 1 of the study ($n = 13$ F2F, $n = 14$ SCOLP). Of the 14 teachers who did not complete the PD or teach the unit in year 1, nine indicated they were assigned to teach grade levels other than seventh or eighth or assigned to teach different content by their district. Fifteen teachers completed their assigned PD and taught the unit 2 years in a row (see Table 2; $n = 8$ F2F, $n = 7$ SCOLP). Of the eight F2F teachers, two had been teaching less than 5 years, three had been teaching between 5 and 10 years, two had been teaching between 10 and 15 years, and one had been teaching over 15 years. Three of the eight had studied biology prior to teaching middle school science, one had studied chemistry, and four had general science degrees. Of the seven SCOLP teachers, one had been teaching less than 5 years, four had been teaching between 5 and 10 years, one had been teaching between 10 and 15 years, and one had been teaching over 15 years. Three of the seven had studied biology prior to teaching middle school, two had studied chemistry, and two had general science degrees.

The average number of students per teacher was 23 for a total of 504 students across Year 1 and Year 2 ($n = 339$ F2F, $n = 165$ SCOLP). Table 3 summarizes the student sample descriptive statistics of consented participating students ($n = 504$). Across the study, gender categories were roughly equivalent, with slightly more female students (51%). Most students in the study identified as White (77%) with 17% identifying as Hispanic. Most students spoke English as their primary language (89%). Ten percent of students in the study received special education services. Because our study was designed as a randomized study, our results demonstrate a high internal validity (i.e., our effects should be unbiased) and causal inferences may be warranted.

Teachers in each PD condition (F2F, SCOLP) completed their PD setting in the summer and then were instructed to teach the *THSB* for 2 years in a row, with different groups of

Table 2. Description of the study sample: participating teachers across years 1–2.

Variables	F2F ($n = 8$)	SCOLP ($n = 7$)
Teaching Experience		
Less than 5 years	2	1
5–10 years	3	4
10–15 years	2	1
Over 15 years	1	1
College Degree		
Biology	3	3
Chemistry	1	2
General Science	4	2

Table 3. Descriptive statistics of the study sample: students across years 1–2.

Variables (n = 504)	Total	n	F2F (n = 339)	n	Online (n = 165)	n
Male	0.47	237	0.46	156	0.48	79
Female	0.51	257	0.53	180	0.50	83
Other	0.02	10	0.01	3	0.02	3
African	0.02	10	0.02	7	0.03	5
Asian	0.01	5	0.01	3	0.01	2
Hispanic	0.17	86	0.26	88	0.10	17
Native	0.01	5	0.01	3	0.00	0
White	0.77	388	0.68	231	0.84	139
Other	0.02	10	0.02	7	0.02	2
English	0.89	449	0.82	278	0.94	155
Spanish	0.10	50	0.16	54	0.04	7
Other	0.01	5	0.02	7	0.02	3
Special Education	0.10	50	0.15	51	0.06	10
Gifted Education	0.01	5	0.01	3	0.00	0
General Education	0.89	449	0.84	285	0.94	155
7th Grade	0.51	257	0.49	166	0.52	86
8th Grade	0.49	247	0.51	173	0.48	79

students. *THSB* was presented as a unit which consists of 19 lessons taught over a six-week timeframe. Most teachers taught the unit in early spring though two taught the unit in Year 2 in late fall.

Measures

Drawing from our conceptual framework and research questions, we gathered data on teacher content knowledge, teacher self-efficacy, change in teachers' instructional practice, and student content knowledge.

Teacher content knowledge

The Teacher Content Knowledge (TCK) assessment was given to participating teachers prior to completing their assigned PD condition and at the conclusion of teaching the *THSB* unit at the end of Year 2. This assessment was developed by the project team to align with the content of the PD and the *THSB* curriculum and was used to better understand change in content knowledge related to NGSS. The assessment was given online and included 16 multiple-choice questions and one open answer item where teachers responded to how their students would answer the question and why. For example, teachers were given a question from the student content knowledge assessment asking students to consider changes in weight for a piece of bread sealed in a plastic bag for 2 weeks when mold begins to grow. Teachers then provided a scientific explanation for the answer choice they selected (e.g., the bag and contents weigh more, less, or stay the same) and then gave reasons for why their students would respond to each of the multiple-choice items (e.g., weigh more, weigh less, weigh the same).

Teacher self-efficacy

The Teacher Self-Efficacy Survey (TSES; Saxton et al., 2014) was also given to participating teachers prior to completing their assigned PD condition and at the conclusion of teaching the *THSB* unit at the end of Year 2. This survey, delivered online, was designed to support teachers in reflecting on their beliefs about their ability to teach science using the NGSS. The TSES measures the potential change over time in teachers' efficacy for effectively using newly learned evidence-based instructional practices in their science instruction. The instrument uses a nine-point Likert rating (not at all to a great deal) as a self-report of teachers' confidence and efficacy in teaching science across six domains: classroom management, student engagement, instructional practices, culturally responsiveness, artistic innovation, and NGSS practices. Internal consistencies overall for the TSES are noted as good (=.94). The variables in the TSES of particularly interest to our study were as follows: (1) student engagement (motivating students, supporting critical thinking, attending to misconceptions); (2) instructional practices (monitoring understanding, guiding questions, using assessment to drive instruction); and (3) NGSS practices (supporting students in developing investigations, crafting scientific explanations, creating models, evaluating explanations).

Change in instructional practices

Potential change in teacher instructional practice was measured using the Student Perception of Instructional Practice Survey (SPIPS; Saxton et al., 2014). Administered to students online at pre- and post-implementation of the *THSB* curriculum by participating teachers, the SPIPS measure is a common instrument designed to assess three research-based instructional approaches that are impactful on student STEM achievement: (1) centeredness—teachers facilitate active engagement of students in their learning with teachers assuming the role of facilitator and students as active learners; (2) assessment—teachers use frequent formative and summative assessments to facilitate diagnostic teaching to create a culture of assessment for learning; and (3) relevance—teachers implement learning activities that students find to be relevant, important, worthwhile, and connected to their cultural and personal lives outside of the classroom and encourage students to use real-world examples in their thinking. Items on the instrument are ranked on a five-point Likert scale (almost never to very often). Internal validity checks across the three domains indicate good internal consistency (0.66). Data collected on the SPIPS included two cycles as participating teachers taught the *THSB* unit to two different groups of students in Year 1 and in Year 2.

Student content knowledge

The Student Content Knowledge (SCK; Herrmann-Abell et al., 2016) assessment was given online to students of participating teachers at pre- and post-*THSB* unit in both Year 1 and Year 2 of the study. This assessment is part of the *THSB* unit curriculum and consists of 21 multiple-choice items and four constructed response items related to content within the *THSB* unit. Using Rasch analysis, item separation reliability was established at 0.96. For the constructed response items, students provided an explanation for their selected multiple-choice item response. For example, students are asked where trees get most of their mass used to make wood. Multiple-choice responses include: (1) from the soil, (2) from oxygen in the air, (3) from carbon dioxide in the air, or (4) from minerals in the soil. In the connected open response item, students are asked to describe the process required to explain their answer choice.

Data analysis procedures

Based on our conceptual framework and research questions, we hypothesized that growth in teacher content knowledge would lead to growth in teacher self-efficacy which would lead to change in teacher instructional practices thus leading to growth in student content knowledge related to the *THSB* unit. We also hypothesized no difference between groups based on their assigned condition (F2F, SCOLP). Thus, our analysis first examined pre- and post-changes for each measure within groups (e.g., change over time by condition) and then compared pre- and post-measures for differences between groups by condition. For continuous outcomes related to teacher level data (e.g., TCK, TSES, SPIPS) and student level data (e.g., SCK), we analyzed within-group data to examine growth using paired samples t-tests of significance and across-group data to examine difference by condition using independent samples t-tests of significance. Difference between groups was determined using Welch's t-test, which is an unequal variances t-test to determine if two samples have unequal variances, particularly when sample sizes may vary (Glass, 1966). To ensure data could effectively be analyzed using an independent sample t-test, the following assumptions were assured: each group was drawn from a normally distributed population with common variance, and all samples were drawn independently of each other. Conducting a Shapiro–Wilk test of normality (Singer & Willett, 2003), each group mean by category (e.g., condition, gender, race/ethnicity, language, special education, grade) was not statistically significantly different from normal indicating they passed the test of normality. While differences within groups existed based on demographics (e.g., grade level), group size differences across PD groups did not reveal variances when conducting a Levene's test of homogeneity (Glass, 1966).

Results

Teacher outcomes

Our first research question examined changes over time in teacher content knowledge, self-efficacy, and instructional practice following participation in the PD condition (F2F, SCOLP) and teaching the *THSB* unit for two consecutive years.

Overall change in teacher content knowledge, measured using the TCK assessment, was the first of three teacher outcome variables of interest for our first research question. Each teacher group (F2F, SCOLP) demonstrated growth from pre- to post-study on the TCK assessment (see Table 4). From pre- to post-study, there was a statistically significant difference for teachers in the F2F condition ($t_7 = 1.17$, $p < .01$) and SCOLP condition ($t_6 = 2.00$, $p < .01$). Pre-study scores were analyzed to ensure there were no differences by condition group prior to conducting post-study analysis. Assuring no difference in pre-study scores by condition, findings indicated no statistically significant difference between condition at post-study on the TCK ($t_{14} = 0.53$, $p = .94$; see Table 5).

Overall change in teacher self-efficacy, measured using the TSES instrument (Saxton et al., 2014), was the second teacher outcome variable of interest in our first research question. The TSES examines change across three domains: student engagement, instructional practices, and NGSS practices. Within-group differences in student engagement from pre- to post study showed an increase for both F2F ($t_7 = 0.42$, $p = .18$) and SCOLP teachers ($t_6 = 0.34$; $p = .11$) which were not statistically significant (see Table 4). Within-group

Table 4. Teacher variables—within-group difference.

	Pre-Study		Post-Study		
	M	*SD*	*M*	*SD*	Difference
Teacher Content Knowledge (TCK)					
F2F	13.50	2.35	14.67	1.86	1.17*
SCOLP	12.14	2.85	14.14	2.12	2.00*
Teacher Self-Efficacy (TSES)					
F2F Student Engagement	6.12	0.42	6.54	0.44	0.42
SCOLP Student Engagement	5.54	0.79	5.88	0.91	0.34
F2F Instructional Practices	6.30	0.63	6.94	0.54	0.64*
SCOLP Instructional Practices	5.50	0.85	5.87	0.91	0.37
F2F NGSS	6.31	1.06	7.27	0.99	0.96*
SCOLP NGSS	5.39	1.33	6.51	1.24	1.12*
Instructional Practices (SPIPS): Year 1					
F2F Centeredness	3.78	0.56	3.83	0.57	0.06
SCOLP Centeredness	3.65	0.61	3.78	0.61	0.14
F2F Relevance	3.66	0.66	3.70	0.71	0.04
SCOLP Relevance	3.41	0.83	3.50	0.82	0.08
F2F Assessment	4.11	0.59	4.14	0.57	0.03
SCOLP Assessment	4.01	0.59	4.02	0.63	0.01
Instructional Practices (SPIPS): Year 2					
F2F Centeredness	3.61	0.59	3.78	0.60	0.17*
SCOLP Centeredness	3.27	0.59	3.62	0.50	0.35*
F2F Relevance	3.46	0.73	3.46	0.85	0.00
SCOLP Relevance	3.26	0.67	3.26	0.74	0.00
F2F Assessment	3.96	.058	4.01	0.65	0.05
SCOLP Assessment	3.92	0.61	3.98	0.56	0.06

*Significant to $p < .01$.

Table 5. Teacher variables—between-group difference at post-study.

	F2F Condition		SCOLP Condition		
	M	*SD*	*M*	*SD*	Difference
Teacher Content Knowledge (TCK)					
Content Knowledge Score	14.67	1.86	14.14	2.12	0.53
Teacher Self-Efficacy (TSES)					
Student Engagement	6.54	0.44	5.88	0.91	0.66
Instructional Practices	6.94	0.54	5.87	0.91	1.07
NGSS	7.27	0.99	6.51	1.24	0.76
Instructional Practices (SPIPS) Year 1					
Centeredness	3.83	0.57	3.78	0.61	0.04
Relevance	3.70	0.71	3.50	0.82	0.20
Assessment	4.14	0.57	4.01	0.63	0.13
Instructional Practices (SPIPS) Year 2					
Centeredness	3.78	0.60	3.62	0.50	0.16
Relevance	3.46	0.85	3.26	0.74	0.20
Assessment	4.01	0.65	3.98	0.56	0.03

*Significant to $p < .01$.

differences for instructional practices from pre- to post study showed an increase for both F2F ($t_7 = 0.64$; $p = .01$), which was statistically significant, and SCOLP teachers ($t_6 = 0.37$; $p = .19$), which was not statistically significant. Within-group differences for NGSS practices from pre- to post-study showed an increase for both F2F ($t_7 = 0.96$; $p < .01$) and SCOLP teachers ($t_6 = 1.12$; $p < .01$), both of which were statistically significant. Assuring no differences between conditions at pre-study, post-study scores between groups in student engagement indicated teachers in the F2F condition scored slightly higher than the SCOLP condition ($t_{13} = 0.66$; $p = .11$) which was not statistically significant (see Table 5). Teachers in the F2F condition scored slightly higher on instructional practices ($t_{13} = 1.07$; $p = .20$) which was not statistically significant (at $p < .01$). Teachers in the F2F condition scored slightly higher on NGSS practices ($t_{13} = 0.76$; $p = .23$) which was not statistically significant (at $p < .01$).

Overall change in teachers' instructional practice, measured using the SPIPS instrument (Saxton et al., 2014), was collected from students for both Year 1 and Year 2 demonstrating their perceptions of their teacher's instruction from before and after teaching the *THSB* unit on three variables: centeredness, relevance, and assessment. For Year 1, within-group differences from pre- to post- study did not show statistically significant growth for the centeredness domain for either F2F ($t_{201} = 0.06$; $p = .13$) or SCOLP teachers ($t_{124} = 0.14$; $p = .10$; see Table 4). Within-group differences from pre- to post-unit did not show statistically significant growth for the relevance domain across years for either F2F ($t_{201} = 0.04$; $p = .43$) or SCOLP teachers ($t_{124} = 0.08$; $p = .22$). Within-group differences from pre- to post-unit did not show statistically significant growth for the assessment domain across years for either F2F ($t_{201} = 0.03$; $p = .48$) or SCOLP teachers ($t_{124} = 0.01$; $p = .94$). For Year 2, within-group differences from pre- to post-study did show statistically significant growth for the centeredness domain for both F2F ($t_{136} = 0.17$; $p < .01$) and SCOLP teachers ($t_{39} = 0.35$; $p < .01$; see Table 4). Within-group differences from pre- to post-unit did not show statistically significant growth for the relevance domain across years for either F2F ($t_{136} = 0.00$; $p = .94$) or SCOLP teachers ($t_{38} = 0.00$; $p = 1.00$). Within-group differences from pre- to post-unit did not show statistically significant growth for the assessment domain across years for either F2F ($t_{136} = 0.05$; $p = .38$) or SCOLP teachers ($t_{39} = 0.06$; $p = .46$). For both Years 1 and 2, pre-unit scores were analyzed to ensure there were no differences at pre-unit by group and year prior to conducting post-unit analysis. Assuring no differences between conditions at pre-unit across years, post-unit scores across groups in (F2F, SCOLP) by year did not indicate unequal variances using Welch's t-test of independent samples. For Year 1, there were no statistically significant differences from pre- to post-unit between groups for the centeredness domain across years ($t_{326} = 0.04$, $p = .51$), for the relevance domain across years ($t_{326} = 0.20$, $p = .18$), or for the assessment domain across years ($t_{326} = 0.13$, $p = .06$, see Table 5). For Year 2, there were no statistically significant differences from pre- to post-unit between groups for the centeredness domain across years ($t_{176} = 0.16$, $p = .13$), for the relevance domain across years ($t_{176} = 0.20$, $p = .18$), or for the assessment domain across years ($t_{176} = 0.03$, $p = .82$; see Table 5).

Student outcomes

Our second research question examined changes in student content knowledge following their teacher's instruction of the *THSB* unit. Each group (F2F, SCOLP) for each year (Year 1, Year 2) demonstrated statistically significant growth from pre- to post-unit on the SCK assessment (see Figure 3). In Year 1, there was a significant difference in post-unit scores for both F2F students ($t_{201} = 1.80$, $p < .001$) and SCOLP students ($t_{124} = 2.54$, $p < .001$). In Year 2, there was a significant difference in post-unit scores for both F2F students ($t_{136} = 3.22$, $p < .001$) and SCOLP students ($t_{39} = 4.56$, $p < .001$). Pre-unit scores were analyzed to ensure there were no differences at pre-unit by group and year prior to conducting post-unit analysis. Assuring no differences, analysis on post-unit scores

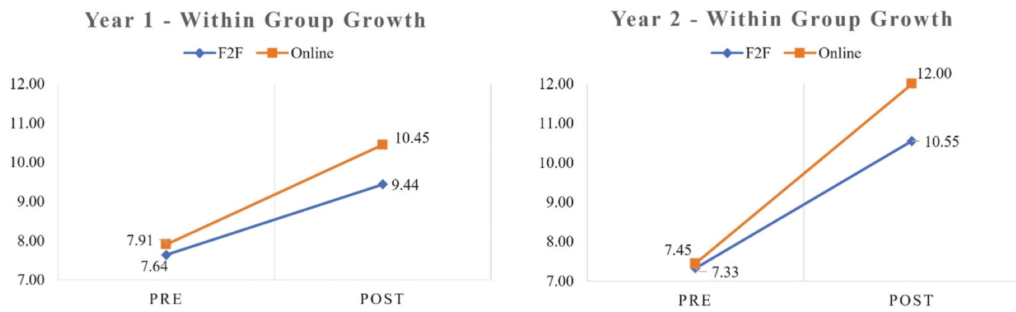

Figure 3. Student content knowledge.

across groups (F2F, SCOLP) by year did not indicate unequal variances using Welch's t-test of independent samples. There were no statistically significant differences between groups for Year 1 on post-unit scores ($t_{326} = 1.01$, $p = .03$) or for Year 2 on post-unit scores ($t_{176} = 1.45$, $p = .03$).

Discussion

Our study examined the development and impact of an online PD, compared with a matched traditional F2F PD, for middle-school science teachers which focused on the use of NGSS practices in an NGSS-aligned curriculum. Overall, our hypothesis was supported: our online PD was as effective as our traditional F2F in promoting growth in teacher content knowledge, growth in teacher self-efficacy related to using NGSS practices, change in instructional practices related to student centeredness, and overall growth in student content knowledge related to the *THSB* curriculum unit. Our findings add to the literature indicating the effectiveness of online PD (Bragg et al., 2021; Tseng & Kuo, 2014; Wasserman & Migdal, 2019), particularly randomized designs that measure the impact on student outcomes (Fishman et al., 2013).

Overall, both teachers and students gained in content learning. Teacher's use of NGSS practices evolved and overall instructional practices changed to be more student-centered. This is a powerful finding in supporting rural teacher PD. Our model, which designed and matched an online PD to a traditional F2F PD, could have the capacity to impact other disciplines (e.g., social studies, math, special education). Research in teacher PD suggests that to enact sustained, systematic change in instructional practice, PD efforts must be "job-embedded" which provides learning within the context of a teacher's day-to-day practice (Althauser, 2015). Our SCOLP PD moved beyond teachers using technology as a tool to communicate (i.e., Zoom, Facetime) or learn content (i.e., Google classroom, Blackboard) to accessing a larger socially connected network of educators engaged in experiencing new learning, sharing problems of practice, and reflecting on their own professional growth for teaching science. The findings from this study have implications to inform the research based on rural school education and essential needs for online PD and learning community building to support fidelity of implementation of professional learning to the classroom.

Our findings indicate that online PD is as effective as F2F in fostering improved science instructional practices. Using this finding, we argue for the importance of designing PD that effectively and economically brings teachers in isolated and diverse rural communities together. As noted above, rural school districts face unique challenges: wide expanse of geographical distance between districts, situated in low-income communities, operating on small budgets. Rural school districts also generally hire only one science teacher at the upper grade level who is responsible for teaching all sciences (e.g., biology, chemistry, Earth sciences, physics). By contrast, urban school districts are more likely to have the resources to connect multiple science teachers within a school building to other science teachers across the district, and bring them together for professional development efforts. Having access to a successful, effective, economic way to provide PD for rural science will impact students and, therefore, be of benefit to rural communities.

Finally, one of the guiding principles of the new vision for science education is promoting equity. All students must have access to high-quality learning opportunities in science (National Research Council, 2015). Effective, sustained, professional learning experiences in science for teachers are needed. This project, and SCOLP in particular, has the potential to impact science teacher quality in rural districts in the sparsely populated part of southwest Kansas, with underrepresented STEM Hispanic populations of greater than 50%. Furthermore, developing high-quality online PD experiences can be transferrable to meet the needs of other rural areas across the US.

Implications

We caution that our findings are based primarily on our PD efforts related to the specific content of our PD, the *THSB* curriculum unit using NGSS practices. While we found statistically significant within-group differences between pre- and post-study measures, some areas showed only practically significant differences. For example, teacher self-efficacy measures (TSES) for two identifiers, student engagement and instructional practices, did show gains between pre- and post-study, but only the F2F group had statistically significant gains in instructional practices (see Table 4). When examining change in instructional practices (SPIPS), we did not see much growth in the relevance or assessment domains across years by group. The *THSB* unit focuses on applying basic chemistry ideas to living and nonliving contexts and covers approximately 6 weeks of instructional time. With more time dedicated to NGSS-aligned curriculum, it is possible that teacher's self-efficacy would improve regarding student engagement and instructional practices and students would see more significant differences in their teachers' instruction related to relevance and assessment.

We purposefully designed our online PD, SCOLP, to utilize evidence-based practices in virtual learning (Blitz, 2013; Bragg et al., 2021; Deterding et al., 2011; Ivanova & Popova, 2011; Tseng & Kuo, 2014). However, we have yet to examine what features of SCOLP were most effective. A vital next step in our work on this project will be an analysis of the features of our matched PDs (see Table 1) that were most effective in impacting positive teacher and student gains. Further research into how our participants engaged with each other and with the features of SCOLP as designed using Stream LXP technology is warranted to propel the field of online PD and professional learning forward.

Finally, our work focused specifically on supporting middle-school science teachers through an introduction to using NGSS practices within an NGSS aligned curriculum. More PD efforts must focus on sustaining those practices, but also addressing student misconceptions through formative and summative assessments. Teachers in rural areas also could use more PD support in place-based instruction that makes science meaningful and relevant, building on rural students' unique knowledge and skills. We are optimistic about the future of online PD and professional learning, specifically for rural areas. Online PD is effective and economical and holds the potential to bridge rural school districts with each other and with PD providers.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Science Foundation [1908937].

ORCID

Brooke A. Moore http://orcid.org/0000-0001-9764-3675

References

Althauser, K. (2015). Job-embedded professional development: Its impact on teacher self-efficacy and student performance. *Teacher Development*, *19*(2), 210–225. https://doi.org/10.1080/13664530.2015.1011346

Barrett, N., Cowen, J., Toma, E., & Troske, S. (2015). Working with what they have: Professional development as a reform strategy in rural schools. *Journal of Research in Rural Education*, *30*(10), 1–18. https://jrre.psu.edu/sites/default/files/2019-08/30-10.pdf

Blitz, C. L. (2013). *Can online learning communities achieve the goals of traditional professional learning communities? What the literature says. (REL 2013-003)*. US Department do Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic. http://ies.ed.gov/ncee/edlabs

Bragg, L. A., Walsh, C., & Heyeres, M. (2021). Successful design and delivery of online professional development for teachers: A systematic review of the literature. *Computers & Education*, *166*, 104158. https://doi.org/10.1016/j.compedu.2021.104158

Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining "gamification." In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*. https://doi.org/10.1145/2181037.2181040

Erickson, A. S. G., Noonan, P. M., & McCall, Z. (2012). Effectiveness of online professional development for rural special educators. *Rural Special Education Quarterly*, *31*(1), 22–32. https://doi.org/10.1177/875687051203100104

Fishman, B., Konstantopoulos, S., Kubitskey, B. W., Vath, R., Park, G., Johnson, H., & Edelson, D. C. (2013). Comparing the impact of online and face-to-face professional development in the context of curriculum implementation. *Journal of Teacher Education*, *64*(5), 426–438. https://doi.org/10.1177/0022487113494413

Glass, G. V. (1966). Testing homogeneity of variances. *American Educational Research Journal*, *3*(3), 187–190. https://doi.org/10.3102/00028312003003187

Herrmann-Abell, C. F., Koppal, M., Roseman, J. E., & Sevian, H. (2016). Toward high school biology: Helping middle school students understand chemical reactions and conservation of mass in nonliving and living systems. *CBE—Life Sciences Education*, *15*(4), 1–21. https://doi.org/10.1187/cbe.16-03-0112

Hjalmarson, M. A., Holincheck, N., Baker, C. T., & Galanti, T. M. (2020). Learning models and modeling across the STEM disciplines. In C. C. Johnson, M. J. Mohr-Schroeder, T. J. Moore, & L. D. English (Eds.), *Handbook of research on STEM education* (pp. 223–233). Routledge.

Hunt-Barron, S., Tracy, K. N., Howell, E., & Kaminski, R. (2015). Obstacles to enhancing professional development with digital tools in rural landscapes. *Journal of Research in Rural Education*, *30*(2), 1–14.

Ivanova, M., & Popova, A. (2011). Formal and informal learning flows: Cohesion in web 2.0 environment. *International Journal of Information Systems and Social Change*, *2*(1), 1–15. https://doi.org/10.4018/jissc.2011010101

Lantz-Anderson, A., Lundin, M., & Selwyn, N. (2011). Twenty years of online teacher communities: A systematic review of formally-organized and informally-developed professional learning groups. *Teaching and Teacher Education*, *75*, 302–315. https://doi.org/10.1016/j.tate.2018.07.008

Lavalley, M. (2018). *Out of the loop: Rural schools are largely left out of research and policy discussions, exacerbating poverty, inequity, and isolation*. Center for Public Education. https://eric.ed.gov/?id=ED608842

Margot, K. C., & Kettler, T. (2019). Teachers' perceptions of STEM integration and education: A systematic review. *International Journal of STEM Education*, *6*(1), 1–16. https://doi.org/10.1186/s40594-018-0151-2

Moser, K. M., & Wei, T. (2023). Professional development in collaborative online spaces: Supporting rural language teachers in a post-pandemic era. *The New Educator*, *19*(1), 1–32. https://doi.org/10.1080/1547688X.2023.2174279

National Center for Education Statistics. (2019). *Nation's report card*. National Assessment of Educational Progress.

National Research Council. (2015). *Guide to implementing the next generation science standards*. The National Academies Press. https://nces.ed.gov/nationsreportcard/

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. The National Academies Press. https://nap.nationalacademies.org/catalog/18290/next-generation-science-standards-for-states-by-states

Powell, C. G., & Bodur, Y. (2010). Teachers' perceptions of an online professional development experience: Implications for a design and implementation framework. *Teaching and Teacher Education*, *77*, 19–30. https://doi.org/10.1016/j.tate.2018.09.004

Reiser, B. J. (2013, September). What professional development strategies are needed for successful implementation of the next generation science standards? In *Invitational Research Symposium on Science Assessment*, Washington, DC.

Saxton, E., Burns, R., Holveck, S., Kelley, S., Prince, D., Rigelman, N., & Skinner, E. A. (2014). A common measurement system for K-12 STEM education: Adopting an educational evaluation methodology that elevates theoretical foundations and systems thinking. *Studies in Educational Evaluation*, *40*, 18–35. https://doi.org/10.1016/j.stueduc.2013.11.005

Showalter, D., Hartman, S. L., Johnson, J., & Klein, B. (2019). *Why rural matters 2018-2019: The time is now. A report of the rural school and community trust*. Rural School and Community Trust. https://eric.ed.gov/?id=ED604580

Siemens, G. (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology & Distance Learning*, *2*(1), 3–10.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford University Press.

Tseng, F. C., & Kuo, F. Y. (2014). A study of social participation and knowledge sharing in the teachers' online professional community of practice. *Computers & Education*, *72*, 37–47. https://doi.org/10.1016/j.compedu.2013.10.005

U.S. Department of Agriculture. (2019). *Rural poverty and well-being*. https://www.ers.usda.gov/topics/rural-economy-population/rural-poverty-well-being/#demographics

Wasserman, E., & Migdal, R. (2019). Professional development: Teachers attitudes in online and traditional training course. *Online Learning*, *23*(1), 132–143. https://doi.org/10.24059/olj.v23i1.1299

Williams, D. (2012). Supporting rural teachers. *Principal*, *26*, 29. https://eric.ed.gov/?id=EJ992661