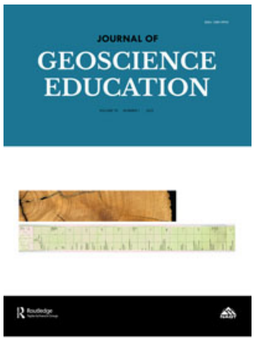

Routledge
Taylor & Francis Group

Journal of Geoscience Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ujge20

Impact of Student Produced Audio Narrative (SPAN) assignments on students' perceptions and attitudes toward science in introductory geoscience courses

E. R. Kraal, G. Sirrakos, L. Guertin, A. Epstein & G. Simmens

JOURNAL OF GEOSCIENCE EDUCATION
https://doi.org/10.1080/10899995.2021.1969863

Impact of Student Produced Audio Narrative (SPAN) assignments on students' perceptions and attitudes toward science in introductory geoscience courses

E. R. Kraal[a], G. Sirrakos[b], L. Guertin[c], A. Epstein[d], and G. Simmens[a]

[a]Department of Physical Sciences, Kutztown University, Kutztown, Pennsylvania, USA; [b]Department of Secondary Education, Kutztown University, Kutztown, Pennsylvania, USA; [c]Earth Sciences, Penn State Brandywine, Media, Pennsylvania, USA; [d]Terrascope Program, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

ABSTRACT

An important challenge for the geoscience community is developing engaging, accessible, and effective experiences within introductory courses, which are often gateways to geoscience majors. This study evaluates low-barrier-to-entry and flexible assignments focused on a pedagogical innovation: faculty replaced one of their usual course components (research paper, lab, etc.) with a Student-Produced Audio Narrative (SPAN) assignment. SPAN assignments require students to engage with geoscience content by telling a scientific story using simple audio-recording and production techniques. The hypothesis is that SPAN assignments will increase students' personal connection to geoscience course content. The pilot study included 8 faculty and 693 students across a range of institution types, course structures, class sizes, and content topics during the control and implementation semesters. The study evaluated student responses to SPAN assignments both quantitatively, using a pre/post survey, and qualitatively, using semi-structured interviews. Survey results show that students experienced positive changes in the categories of personal relevance, sense of curricular innovation, and future intentions to study science. Interview results indicate that much of the increased senses of innovation and personal relevance came from the creativity and choice the students experienced during the SPAN assignment. Taken together, these results indicate that SPAN assignments are innovative to students and effective pedagogical tools that can positively change students' perceptions of their learning environment and attitudes toward science.

ARTICLE HISTORY
Received 5 October 2020
Revised 9 August 2021
Accepted 13 August 2021

KEYWORDS
Narrative; learning environment; curricular innovation; introductory courses; audio

Purpose

Introductory geoscience courses play a key role in building a more diverse and larger geoscience workforce as well as a science-literate public. Exit surveys reveal that nearly one-third of geoscience majors choose their major as a result of an early experience, such as an introductory course, in either community college or the first undergraduate year (Wilson, 2016). These students enroll in an introductory geoscience course for reasons other than a major, such as general education requirements, but then leave on a path to the geosciences. Introductory courses serving non-STEM majors are valuable recruitment opportunities for increasing the number and diversity of geoscience majors. However, there is a large and well-documented disparity between the number of students enrolled in introductory courses and the small percentage who become geoscience majors (Martinez & Baker, 2006). Therefore, introductory courses must employ new and different pedagogical techniques to help close this gap.

Reaching introductory students is critical but difficult (e.g., Egger, 2019a). Geoscience faculty teaching introductory courses, particularly at two-year colleges (2YCs), face challenges such as: difficulty providing research, field, and laboratory experiences; discipline isolation; and students' limited exposure to geoscience careers and opportunities (Wilson, 2014). Experiences and activities focused on introductory teaching have been the focus of many research studies, programs, and discipline-wide efforts (such as SERC's "The Starting Point," "On the Cutting Edge," etc.). Yet field courses, place-based experiences and undergraduate research, while effective, can be difficult to implement in many general education, introductory-level courses (National Academies of Sciences, Engineering, and Medicine, 2015, 2017). The reality is that most general education, introductory courses involve large lecture halls, few or no teaching assistants and little teaching budget. Students are often juggling multiple jobs, are first-time college students, and/or are taking care of a family (Holm Adamec & Asher, 2013). Weekend-long field trips and after-hours undergraduate research are not feasible options for a large number of these students. However, given the established importance of early

CONTACT E. R. Kraal kraal@kutztown.edu Department of Physical Sciences, Kutztown University, 15200 Kutztown Rd., Kutztown, PA 19530, USA.

experiences and courses for students who might choose a geoscience major and career, it is important to reach these students (Egger, 2019a).

Learning goal

The research presented here addresses designing and evaluating student experiences within these critical introductory courses. We evaluated a flexible, adaptable assignment approach called Student-Produced Audio Narrative (SPAN) for instructors to use in introductory level geoscience courses. For the purposes of this study, we define geosciences according to the AGI Workforce Report, which includes geology, hydrology, planetary science, marine science/oceanography, atmospheric and space sciences, climate science, geochemistry, petrology, paleontology, environmental sciences and related fields (Wilson, 2016). In SPAN assignments, students engage with the geosciences by telling a scientific story using simple audio recording and production techniques. Using qualitative and quantitative methods, we examined the impact of SPAN assignments in terms of students' attitudes toward science and their perceptions of the classroom learning environment.

This research was inspired by the experiences of authors Epstein, Guertin, and Kraal, who found that including audio-production projects in introductory-level classes engaged students in new and powerful ways, and reached students who might otherwise never see the relevance and importance of the subject matter (Epstein et al., 2010; Guertin, 2012; Kraal & Regensburger, 2013). These observations led to an NSF-IUSE pilot grant to formally explore educational outcomes of the SPAN approach.

Literature context

To address the challenges of teaching introductory courses, we must establish inclusive, adaptable pedagogies that work specifically within the structures of general education courses. SPAN assignments achieve this through several approaches.

Inclusive pedagogy

Addressing equitable instruction involves incorporating inclusive pedagogical approaches. One approach is constructivism, in which students have opportunities to make sense of the world in relation to their emerging knowledge, thus emphasizing the personal relevance of the content (Dorman & Adams, 2004). This is particularly important in science classrooms where, all too often, science is taught as a collection of facts and patterns, devoid of any connectedness to students (Emdin, 2010; Keat & Urry, 1982). SPAN assignments provide opportunities for students to link new ideas to familiar ideas, enabling them to internally construct understanding. Further, SPAN creates an inclusive learning environment that acknowledges, values, and utilizes students' unique and diverse cultural resources as tools to learn science, resulting in a shift away from traditional notions of assessment and learning. Student narratives may have a role in intersectionality exploration, in which students select a topic that is socially relevant to themselves and their lived identities (Nuñez, Rivera, & Hallmark, 2020).

This study examines scientific storytelling through an audio-narrative approach, rather than through written essays or video production. Audio is viewed as an inclusive pedagogy for student learning (see Kleege & Wallin, 2015; Schmidt, 2013). Students who are challenged with long commutes to campus, who spend time exercising, or who engage in a variety of other activities may have the opportunity to listen to audio for academic or nonacademic purposes. Audio creation provides students ownership of the audio design and format, the methods of expression and communication, all within the assignment guidelines (see the National Center on Universal Design for Learning, http://udlguidelines.cast.org/).

Accessible audio

Audio has a number of unique qualities that make it well-suited in introductory courses. First, there is a low barrier to entry in terms of cost and training. Basic audio can be recorded using only mobile phones. Free software for audio editing is widely available and accessible on common platforms. Second, audio files are small enough to be transferred and manipulated easily. These considerations are particularly relevant where interaction often occurs through online course management systems. Third, audio assignments are novel to students. The majority of assessment in geoscience courses is from exams (Egger, 2019b). While most students have produced papers, written laboratory reports, given presentations, or created websites, very few of them may have engaged in audio storytelling, so they are not bound by prior experience and expectations. Finally, today's students surround themselves with audio experiences. As of 2019, nearly 51% of Americans older than 12 had listened to an audio podcast within the past month, and that number increases annually (Pew Research Center, 2019).

Student produced

The majority of studies on pedagogy and audio production focus on the dissemination of content *to* students, not the creation of content *by* students (e.g., Drew 2017). Few studies, like the results presented in this paper, focus on the outcomes of audio creation by students. Note, some studies use the terminology of "podcast" to denote what in this study we term an "audio piece." Technically, the term "podcast" refers to a particular mode of distribution, which is not part of the SPAN assignments. Therefore, we use "audio piece," "audio assignment," or "audio narrative" to denote students modeling this assignment with the intent of student learning versus professional media products. For example, in Lee et al. (2008), investigators found that "learner generated" audio stories facilitated collaborative knowledge building and disciplinary knowledge in an introductory-level information technology course. Pegrum et al. (2015) used a controlled study implementing "creative podcasts" created by students in an introductory chemistry class. Among their results were that the creation of audio assignments by students was more influential than listening to audio created by others, and that this technique shows promise for large, introductory classes. Some other research has similar results, that student-created audio production is an excellent medium for establishing connections in isolated college populations (Lee & Chan, 2007), and audio production can inspire students from many backgrounds to engage in their education (Campbell, 2005).

Further, the argument for instructional practices that use multimodal representations of content has been well established (Dhingra, 2008; Lundby, 2008; O'Neill & Calabrese-Barton, 2005; Robin, 2006, 2008). These studies assert that student-produced digital media cultivates a sense of ownership over the content learned, resulting in increased engagement with the learning environment. Audio assignments, such as SPAN assignments, may have the capacity to encourage students' critical thinking and discourse as society moves to new media and communication technologies (Deal, 2007).

Narrative

The audio pieces created by students in SPAN assignments have another important component: the focus on narrative, or storytelling. Dahlstrom (2014) described science narrative or storytelling in this way: *Narratives follow a particular structure that describes the cause-and-effect relationships between events that take place over a particular time period that impact particular characters…[the] triumvirate of causality, temporality, and character represents a fairly standard definition of narrative communication.*

The role of cause-and-effect (time and impact) and relationships are important. Narratives contrast with other types of communications, such as lists, descriptions, or definitions. Ultimately, the key to developing a narrative is establishing connections between events and information. Scientific stories have been examined from a pedagogical standpoint, where scientific narratives were developed as instructional tools for delivering STEM content (e.g., Clough, 2011; Klassen, 2009). Others have shown that narrative can help scientists connect and communicate with the public (e.g., Dahlstrom, 2014; Gardiner, 2018).

The connections involved in narrative or storytelling may be especially critical, given that researchers have recently found that personal interest is the main reason students select a geoscience major (Sexton et al., 2018). Developing a scientific narrative or story requires students to move beyond a repetition of facts and to look for connections. These connections could be personal (such as a childhood experience or an event in their community) and/or across disciplines (such as linking to political or economic considerations). SPAN assignments take authentic root in the students' interests and enthusiasms. Through these assignments, introductory-level students can develop and explore their interest in (geo)science and, therefore, they serve as valuable mechanisms in building interest among prospective geoscience-majoring students.

Study population and setting

The purpose of this pilot study was to examine the student response to completing a SPAN assignment as a part of their introductory geoscience course, with faculty teaching in diverse settings who had not previously used this approach. This pilot study took place over 4 academic semesters starting in Fall 2017 and concluding in Spring 2019. It included one semester of training pilot faculty (Fall 2017) and three implementation semesters (Spring 2018, Fall 2018, Spring 2019). Two of the NSF PI team (coauthors Kraal and Guertin) also continued to implemented SPAN in their courses with data collected from their students. Although these instructors were not part of the participant faculty group, these data were included to increase the sample size.

Participant faculty

The pilot program trained a cohort of six faculty within a regional area, all teaching at least one introductory geoscience course. Faculty were drawn from both traditional geoscience programs and multidisciplinary programs such as environmental science. The faculty represented a wide range of public and private institutions including urban, suburban, and rural; community colleges, R1 university, and regional state universities. Faculty classifications included temporary part-time and full-time as well as pre- and post-tenure. All courses were taught face-to-face, though all the faculty did use course management systems to facilitate delivery of the SPAN assignment. One goal for the pilot study was to observe SPAN implementation in a variety of course situations that corresponded to typical introductory courses; see Table 1.

In the fall and winter of the 2017–18 academic year, the participant faculty received three days of training to introduce the SPAN approach and to assist faculty in developing their own SPAN assignments for use in their specific courses. Each faculty member created or modified an assignment within a course to include the characteristics of a SPAN assignment (described above). Following the face-to-face training period, the faculty were supported with online, virtual meetings about once a month during the semester with more continuous mentoring from the PI team. During the course of the study, faculty were free to modify their SPAN assignments.

Student information

There were a total of 693 student participants in the study. Of this total, 146 students were part of the control group (Fall 2017). The remaining 547 students were part of the implementation phase of the study. Over the course of the

Table 1. Summary of SPAN assignments developed by partner faculty used in the pilot study. All courses were introductory level, non-STEM major courses. This table does not include PI's course information from coauthors Kraal and Guertin (see link in text for an example of a full SPAN assignment).

Subject area	Earth science	Physical geology	Environmental science	Space science	Environmental science	Planetary science
Class type, Size	Lecture, ~25 students	Lecture with lab, <20 students	Lecture, 50–100 students	Lecture with lab, ~25 students	Lecture, ~30 students	Lecture, ~40 students
% of Total course grade	20%	35% (audio + paper)	25%	25% (audio + lab report + pre-lab + tutorial assignments)	30%	25% (two audio productions)
Student team size	Group	Individual	Variable	Variable	Individual, supported by small peer group	Individual or pairs
Assignment summary	Two-minute audio with bibliography/sources. Students' choice of format (e.g., voicemail message, newscast, story from point of view of natural phenomenon, etc.), students' choice of topic	Three-minute audio and 1200-word research paper. Students' choice of format (e.g., journal, poetry, rap, radio play, interview, etc.), students' choice of topic	Simulated call-in radio show. Topic chosen by instructor (same for whole class)	Two- to three-minute audio drama accompanying one-page lab report (audio simulates narrative from fictitious commercial lab). Topic chosen by instructor (same for whole class)	Audio piece based on news article. Format open (e.g., commentary, audio postcard, etc.). Students' choice of topic	Two 4–6-minute pieces. Open format, emphasizing creativity (e.g., song, radio drama, travel narrative, voicemail message, etc.). Audio effects/background sound/etc. required; specified scientific information required. First project: meteorites. Second project: camping trip to Mars
Project stages	Topic proposal; outline; two script drafts; final production	Topic; research paper; script; final production	Scaffolded assignments building toward final product	Pre-lab activities; audio-tutorial assignments; lab report; final production	Library research; topic definition and presentation; raw audio + transcript; script; final production	Description of topic and audio format; script and sample of sound; final production
Scaffolding/Support provided	Audio narratives played and discussed in class; Audacity-introduction lab; Audacity-practice lab; Audacity handouts. Detailed grading rubric	Scaffolded assignments early in semester (narrative-writing, recording, listening, topics, sources, sound-layering); writing-center tutors; production lab session. Detailed grading rubric	Scaffolded assignments throughout semester (scientific concepts, development of viewpoint, facts/ sources, collaboration, scripts). Detailed grading rubric	Staffed academic computing center; links to video tutorials on audio-editing software; short assignments (narration with music, interview, script) on different topic as part of pre-lab activities. Detailed grading rubric	Library/source instruction; audio practice sessions; technology classroom; peer evaluation. Detailed grading rubric	"Practice with technology" mini-assignment; "practice editing" mini-assignment; links to scientific papers and other sources; lectures on topics. Step-by-step instructions for common audio-editing tasks; Audacity documentation and links to tutorials. Detailed grading rubric

three academic semesters of implementation, 547 students in 25 different courses completed post-surveys that included demographic information giving insights into aspects of the student population. Table 2 summarizes student characteristics.

Materials and implementation

SPAN assignments

SPAN assignments utilize audio storytelling to bring together compelling, student-centered, active pedagogical components that can be implemented in a variety of introductory courses. These components are grounded in previous research outlined above and combined here in a unique, inclusive, and flexible pedagogy. For the purpose of this study, a SPAN assignment must:

- be narrative in nature, focused on storytelling rather than reporting or listing information
- demonstrate students' understanding of scientific concepts in the course as appropriate to the assignment and course objectives
- ultimately be submitted as an audio file in which a student or group of students has recorded and edited audio sounds, such as their own voices, perhaps ambient sound, etc.

The focus of a SPAN assignment is defined by the instructor and can be broadly interpreted (e.g., anything that interests students about the geosciences) or content-focused (e.g., about students' interaction with the water cycle); see Table 1. Assignments were not designed or evaluated based on audio quality—for example, production of "broadcast-quality" pieces. The focus in this study was on the *process* of student audio-narrative creation not the final *product*.

Table 2. Distribution of student population characteristics. Total number of completed student post-surveys (547) across three semesters. College credits was used as an approximation for year in college.

Category	Distinction	% of students
Major	STEM Major	12.8
	Non-STEM	87.2
Gender	Female	44.6
	Male	54.9
	Other/Prefer not to identify	0.5
Race/Ethnicity	Non-Hispanic white	58.1
	Hispanic or Latino	14.5
	Asian or Asian American	12.4
	Black or African American	11.9
	Mixed Race	1.4
	American Indian, Alaskan Native, Hawaiian or other Pacific Islander, Middle Eastern, or n/a – each comprising less than 1%	1.7
College credits	0–30 (First year)	30.6
	30–16 (Sophomore)	35.9
	60–90 (Junior)	17.7
	More than 90 (Senior)	15.0
	n/a	0.7

Developing SPAN assignments

For readers interested in implementing SPAN assignments in their own classes, we briefly describe the process pilot faculty underwent as a part of the research program in developing their assignments, along with a summary of their assignments (Table 1).

Three day-long training workshops were distributed across a semester and combined with asynchronous activities during the semester. First we introduced the contemporary world of audio production and the wide range of possible audio approaches. Sharing multiple, varied examples of audio production (short documentaries, public-service announcements, audio dramas, profiles, personal reflections, etc.), we encouraged participants to listen for particular elements and to reflect on the storytelling techniques embodied in these audio pieces. We also discussed the nature of narrative and how narrative/storytelling differs from simply reporting or recounting facts (Dahlstrom, 2014) by reflecting on the ways in which the pieces use narrative/storytelling.

Next we introduced pilot faculty to audio recording and editing, using their mobile phones as microphones and the free, cross-platform audio-editing software Audacity (https://www.audacityteam.org/) for production. We provided some basic technical support to some pilot faculty. However, Audacity is a freeware program with many Audacity tutorials available online. We also encouraged faculty to tap into other freely-available software on campus for students to use, such as GarageBand and Adobe Audition (not generally free, but available via site license at some institutions), as local on-campus technical support should be able to address any site-specific issues.

Next, the pilot faculty selected an assignment they ordinarily give in their regular classes to transition to a SPAN assignment. An important part of this process involved thinking through the learning goals inherent in each assignment. Support for this type of analysis is found in curricular and pedagogical materials, such those linked on the SERC Site Guide for Designing Courses (https://serc.carleton.edu/36759). Pilot faculty determined which learning objectives might best be accomplished using a SPAN assignment. They then filled out worksheets defining learning goals for the audio assignment, the scaffolding support students would need (library, writing center, etc.), and other elements. We then worked with them to develop and refine SPAN assignments to distribute in their classes. Pilot faculty had the discretion to determine the scope of the assignment as appropriate for their course; assignments encompassed a range from an extended two-week lab assignment to a semester-long project. Some instructors used groups; others used individual projects. Table 1 summarizes assignments developed by pilot faculty in this study.

To provide concrete examples for those interested in developing their own SPAN assignments, two example full SPAN assignments authored by coauthor Guertin can be found in NAGT's Teach the Earth portal, Rally Speeches for Coastal Optimism (https://serc.carleton.edu/240117) and FutureEarthCast (https://serc.carleton.edu/243300)

Before the pilot faculty implemented their first SPAN assignment, we discussed some possible items to be aware of. First, was the technology available to students on campus – e.g., does audio recording/editing software exist on campus laboratory computers? Is this something faculty can request to be downloaded/installed? We did have one faculty participant at an institution where students were not allowed to plug in microphones to any campus computers, and we addressed the opportunities and apps available for recording on mobile devices such as cell phones. We also discussed working with students that may be deaf or hard-of-hearing, or may have a speech impediment. We again emphasized that a SPAN assignment emphasizes the process and not the final technical product, and a student could be responsible for developing the script for an audio narrative but have someone else record it. For example, one faculty member had a student with a stutter, and that student wrote the script, had another student read it, and then completed the final editing before submission. We emphasize that in all cases, ultimately the focus of the assignment was on the student process of creation, not the final product.

Research methods

Overall design

In this paper, we report results for one of our research questions, as part of a larger NSF-funded research study. Specifically, we sought to explore and understand the impact of SPAN assignments in terms of students' attitudes toward science and their perceptions of the classroom learning environment. In line with the research hypothesis noted above, this study did not evaluate SPAN assignments for content learning or for audio production quality. As recommended by Tobin and Fraser (1998), we chose to use a mixed-methods approach involving a questionnaire to collect quantitative data and semi-structured interviews to collect qualitative data.

Data sources & collection

Collecting quantitative and qualitative data increased the reliability of our results (e.g., Fraser & Tobin, 1991) and provided us with a more holistic and robust understanding of the impact of SPAN assignments on students. We also wanted to accurately capture the impact of SPAN assignments on how students developed and/or changed specific attitudes toward science as well as their perceptions of key aspects of the learning environment, particularly those that we preliminarily identified as important in science classrooms. Further, our recognition of the limitation of a single type of data source (i.e., a student questionnaire) as possibly containing too much error to be reliable by itself (Fraser & Tobin, 1991) prompted us to include semi-structured interviews as additional data.

Quantitative data collection

For this research study, we collected quantitative data using the Questionnaire Assessing Connections to Science (QuACS). Sirrakos and Fraser (2019) developed and field-tested the QuACS to "evaluate educational innovations that involve place-based learning, scientific storytelling or narratives" in terms of students' perceptions of the classroom learning environment and their attitudes toward science. As part of the instrument's development, Sirrakos and Fraser (2019) demonstrated the QuACS' "sound factorial validity and internal consistency reliability."

The QuACS consists of 47 items across six scales (Table 3). Four of these scales are dedicated to assessing students' attitudes toward science: *Future Intentions to Study Science*, *Science Self-Efficacy*, *Scientific Storytelling*, and *Place-based Learning*. The remaining two scales assess students' perceptions of specific dimensions of their learning environment, namely *Personal Relevance* and *Innovation*. Each of the 47 items was scored using a Likert scale of 1, 2, 3, 4, and 5 respectively, for the response choices of Strongly Disagree, Disagree, Neutral, Agree, and Strongly Agree. Table 3 provides an overview of each QuACS scale including its name, a scale description, a sample item, and the number of items associated with the scale (see Sirrakos & Fraser, 2019 for the full questionaire).

QuACS was administered as a pretest and posttest to students during the academic semesters Fall 2017, Spring 2018, Fall 2018, and Spring 2019. The course instructor was responsible for administering the QuACS. The research team provided each instructor with a list of directions and a script to read to students. The instructor informed students that their questionnaire responses would remain confidential and not affect their course grade. This information, along with additional directions and a statement of consent, was printed on the cover sheet of each questionnaire. Students completed the pretest at the beginning of the semester, prior to any discussion relating to or any experience with a SPAN assignment. Students completed the posttest at the end of the semester, after having completed their SPAN assignment. This was true for all semesters except Fall 2017, whose students served as a control group, and thus no SPAN assignments were given. Our analyses required us to match pretests and posttests, thus anonymous questionnaires were not possible. Instead, they remained confidential. Due to the pilot study size, all data in all courses were binned for analysis. Survey data for individual course or student or instructor were not analyzed, and faculty did not receive any individualized or grouped course information. Response rate was 55% for matched pre-post surveys. As our data analyses required matched pretests and posttests, there was some level of attrition, for instance, where a student completed a pretest, but not a posttest, or vice versa. These students were not included in our analyses. The research team was not informed by pilot faculty of students refusing to participate.

Qualitative data collection

Semi-structured key informant interviews of students were used to further investigate students' attitudes toward science

Table 3. Structure of the questionnaire assessing connections to science (QuACS).

Scale	Description	Sample item	Number of items
Personal relevance	The extent to which school science connects with students' out-of-school experiences	This course provides me with a better understanding of the world outside school.	7
Innovation	The extent to which the instructor utilizes a variety of new activities, teaching techniques and assignments.	New and different ways of teaching are used in this class.	7
Future intentions to study science	The extent to which students indicate their intentions to study science in the future or pursue a science-related career	I intend to study science in the future.	7
Science self-efficacy	The extent to which students believe that they can be successful in science and scientific communication	I am confident I can do well in this science course.	14
Scientific storytelling	The extent to which students believe that scientific storytelling assists them in making connections to science	Combining scientific information from several sources into a story is an interesting way to learn science.	7
Place-based learning	The extent to which students believe that the local community is a good source of science learning	The local community is a useful resource for learning science.	5

and perceptions of the learning environment in response to SPAN assignments. By utilizing key informants, we had the opportunity to understand the personal perspectives of a small sample of individual students regarding the impact of SPAN assignments that might not necessarily have to come to light through the QuACS alone (Rubin & Rubin, 2005). As suggested by Morgan (1997), the semi-structured format provided the research team with the needed flexibility to begin with a small number of pre-established questions and then ask follow-up questions that made sense in terms of the particular context of the interview and the themes that emerged in real time.

Toward the end of each implementation semester (Spring 2018, Fall 2018, and Spring 2019), faculty participants were asked to identify three students, or key informants, who might be willing to participate in an interview. Each of the identified key informants in each class had a different experience with SPAN: one was identified as enjoying SPAN, another was identified as seemingly apathetic toward SPAN, and another was identified as someone who did not enjoy SPAN. In an attempt to minimize bias during our analyses, the research team intentionally did not monitor which group (i.e., students who enjoyed SPAN versus students who did not seem to enjoy SPAN) individual students interviewed represented. Not every student who was identified by a faculty participant participated in an interview. Over the course of the three implementation semesters, 54 students were identified and nine agreed to participate. The nine interviewees represent a participation rate of approximately 17%.

In line with the semi-structured format, the interviewer used the following set of questions as starting points for each interview:

- How did the SPAN assignment differ from other types of assignments completed in your other courses?
- Did you enjoy completing the SPAN assignment? Why? Why not? Was there a particular aspect of the assignment that you liked best? Least?
- Did you feel that the SPAN assignment made the course more engaging? Why? Why not?
- What did you perceive as the advantages and limitations of using the SPAN assignment as a tool to learn science content?
- When completing the SPAN assignment, how did you attempt to make the science content more engaging and relevant for the listener?
- How would you describe the SPAN assignment to a student who could register for the course next semester?

To ensure that interviews were conducted effectively and consistently, a set of interview protocols was developed as described by Anderson (1998). The approximately 30-minute interviews were conducted generally within two weeks of the course ending and completed using video conferencing software, which also made recording each interview easier. With the capacity to record, the interviewer did not have to take notes during the interview and, instead, could listen intently and ask follow-up questions as appropriate.

To ensure the trustworthiness of our qualitative data, at the conclusion of each interview, the interviewer provided the participant with an oral synopsis of key points shared during the interview to ensure accuracy. Additionally, the transcribed interview data were initially reviewed separately and then analyzed jointly by two members of the research team.

Validity and reliability

Measuring an instrument's validity ensures that what we are measuring appropriately reflects what we expect to measure. Establishing an instrument's reliability ensures that each of the items making up a unique scale reflect a common construct and that there is consistency among the associated items. These types of analyses are important when working with a questionnaire, because they add credibility to the results and conclusions made based on data obtained using this instrument.

For this study, factor analysis of the QuACS was carried out using principal axis factoring with Oblimin rotation and Kaiser normalization to check the structure of the 47-item, six-scale questionnaire. Oblimin rotation was suitable as we anticipated the scales being related, rather than independent. Kaiser normalization facilitated obtaining stable factor solutions across different samples. In contrast to the preliminary validity testing of the instrument (Sirrakos & Fraser, 2019) for which one factor analysis was conducted for the two learning-environment scales and a second factor analysis was conducted for the four attitude scales, all subsequent analyses involved a single factor analysis that included all six QuACS scales. Pretest and posttest analyses were conducted separately for the sample in each of the four semesters (Fall 2017, $N = 146$; Spring 2018, $N = 141$; Fall 2018, $N = 204$; and Spring 2019, $N = 202$), as well as for the combined sample ($N = 693$ including the control and $N = 547$ for study semesters). For each of these factor analyses, the factor solution closely aligned with our instrument's six-scale structure, thus further supporting the instrument's validity. Factor analysis results for the combined sample displayed patterns similar to those for the specific data set each semester.

Table 4 presents factor loadings for the QuACS for the pretest and posttest for the entire sample. We used two criteria when making decisions whether or not to retain an item. In order to retain an item in the QuACS, it had to have a factor loading of (1) at least 0.40 on its own scale and (2) less than 0.40 on all other scales (Tabachnick & Fidell, 2013). Factor analysis of pretest data revealed that all 47 items satisfactorily loaded on the correct scale and not on other scales. Factor analysis of posttest data revealed similar findings, but with some minor discrepancies. For the posttest data, factor analysis results for five of the six scales were consistent with the pretest data. The items associated with *Place-based Learning* did not load on their own scale, but rather loaded with *Personal Relevance*. This factor loading issue only appeared for the posttest for the entire sample, but did not appear during preliminary analyses of the sample for each semester. Additionally, three items on the *Science Self-Efficacy* scale did not adequately load on their own scale, but this is inconsequential because of the large number of items associated with that factor/scale.

Further, as part of assessing the instrument's validity and factor strength, we calculated eigenvalues and percentages of variance for each scale (see bottom of Table 4). The percentages of variance for the QuACS pretest for the total sample ranged between 2.87% to 37.58% and the eigenvalues ranged between 1.35 and 17.66 for different scales. The total proportion of variance accounted for by these 47 items in six scales was 64.54%. Percentages of variance for the posttest for the total sample ranged between 2.65% to 50.16% and the eigenvalues ranged between 1.25 and 23.58 for different scales. The total proportion of variance accounted for by these 44 posttest items in five scales was 70.16%. Overall, these factor-analysis results support the structure and validity of the QuACS.

To determine each scale's internal consistency reliability, we calculated the Cronbach alpha coefficient. Because the alpha coefficient has an upper bound of 1, an alpha coefficient for a particular scale approaching 1 provides greater confidence that the items indeed reflect a common construct (Boslaugh & Watters, 2008). According to Cortina (1993), reliability values over 0.70 are desirable. Cronbach alpha coefficients across the six scales for the whole sample ranged from 0.86 to 0.95 for the pretest and from 0.94 to 0.97 for the posttest. (For individual Cronbach alpha values see bottom Table 4). These results support the QuACS' strong internal consistency reliability.

Quantitative data analysis

For the pretest and posttest QuACS data collected in each semester (Fall 2017, Spring 2018, Fall 2018, and Spring 2019), Table 5 provides the mean and standard deviation for each QuACS scale for each semester. Fall 2017 data (shown in *italics* in Table 5 to facilitate comparison) involve a student control group who did not complete a SPAN assignment. The other three semesters involved experimental groups of students who each completed a SPAN assignment.

In each semester, students' matched pretest-posttest scores for each of the six QuACS scales were analyzed to investigate changes over time. First, a repeated-measures ANOVA was calculated for each semester for each QuACS scale to ascertain the statistical significance of changes between pretest and posttest. Additionally, an effect size (d) was calculated to indicate the magnitude, or educational importance, of the pretest–posttest difference for each scale without being burdened by the limitation of sample size (Cohen, 1992). Calculating effect sizes is important because the presence of a statistically-significant difference between two means, which might be primarily a function of sample size, does not in itself indicate whether the difference is important. Therefore, combining traditional significance testing (ANOVA) with a calculation of effect size provides richer insights. According to Cohen (1992), an effect size of 0.2 is considered "small," of 0.5 is "medium," and of 0.8 and above is "large."

Finally, ANCOVA (F) was used to compare the control group with each of the three experimental groups on posttest scores on each QuACS scale while controlling statistically for corresponding pretest scores. Using the pretest as covariate in ANCOVA is a more sophisticated statistical test than directly comparing pretest-posttest changes for two groups because of the potential unreliability of change scores. ANCOVA results are in the last column of Table 5.

Qualitative data analysis

Analysis of semi-structured student interviews elaborated on several findings from the QuACS and revealed additional student perspectives regarding the impact of SPAN assignments. Each interview was transcribed and then reviewed by two members of the research team to identify common

Table 4. Factor analysis results for QuACS for pretest and posttest for combined sample.

	Factor loadings											
	Personal relevance		Innovation		Future intentions to study science		Self-efficacy		Scientific storytelling		Place-based learning	
Item	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
PR1	0.43	0.45										
PR8	0.52	0.65										
PR14	0.50	0.50										
PR21	0.40	0.58										
PR28	0.40	0.54										
PR34	0.48	0.57										
PR41	0.44	0.46										
INN2			0.70	0.79								
INN9			0.82	0.81								
INN15			0.74	0.74								
INN22			0.71	0.86								
INN29			0.78	0.86								
INN35			0.61	0.63								
INN42			0.84	0.94								
INTEN3					0.85	0.84						
INTEN10					0.58	0.70						
INTEN16					0.86	0.79						
INTEN23					0.73	0.73						
INTEN30					0.83	0.78						
INTEN36					0.91	0.88						
INTEN43					0.75	0.81						
EFF4							0.63	–				
EFF5							0.70	0.60				
EFF11							0.71	0.78				
EFF12							0.77	0.88				
EFF17							0.74	–				
EFF18							0.72	0.76				
EFF24							0.58	0.63				
EFF25							0.77	0.89				
EFF31							0.75	0.60				
EFF32							0.75	0.83				
EFF37							0.73	–				
EFF38							0.71	0.78				
EFF44							0.63	0.54				
EFF45							0.72	0.72				
STORY6									0.60	0.56		
STORY13									0.61	0.61		
STORY19									0.66	0.62		
STORY26									0.71	0.77		
STORY33									0.91	0.81		
STORY39									0.73	0.74		
STORY46									0.75	0.65		
PLACE7	–	0.60									0.63	–
PLACE20	–	0.55									0.76	–
PLACE27	–	0.45									0.71	–
PLACE40	–	0.55									0.62	–
PLACE47	–	0.55									0.61	–
% Variance	3.95	50.16	4.77	8.69	5.01	4.68	37.58	3.98	10.36	2.65	2.87	–
Eigenvalue	1.86	23.58	2.24	4.08	2.36	2.20	17.66	1.87	4.87	1.25	1.35	–
Alpha Rel.	0.87	0.97	0.94	0.94	0.94	0.94	0.95	0.96	0.86	0.94	0.86	–

Principal axis factoring with Oblimin rotation and Kaiser normalization.
N = 693.
Factor loadings smaller than 0.4 omitted from table.

themes. We recognized a theme as an idea specifically repeated multiple times during an interview or across several interviews or something communicated with importance. Some of the identified themes aligned with scales on the QuACS, while other themes captured additional dimensions of the student experience not explicitly defined on the QuACS. Once the team identified preliminary themes, these were reviewed for frequency and organized hierarchically to examine the interconnectedness among themes and to determine primary and secondary themes. Direct quotes from student interviews were extracted and organized to support the thematic analysis. This process occurred at the end of each implementation semester and again at the study's conclusion.

Table 5. Pretest–posttest changes (ANOVA results and effect size) and ANCOVA results for comparison of experimental and control groups on posttest while controlling for pretest for QuACS scales.

Area	Scale	Group	Mean Pre	Mean Post	SD Pre	SD Post	Pre–post change F	Pre–post change Cohen's d	ANCOVA F
Learning environment	Personal relevance	*Fall 2017 Control*	*3.76*	*3.74*	*0.79*	*0.85*	*0.16*	*−0.03*	–
		Spring 2018 Exptl	3.53	3.69	0.77	0.91	7.62**	0.20	2.64
		Fall 2018 Exptl	3.62	3.80	0.70	0.80	15.68**	0.24	2.81
		Spring 2019 Exptl	3.62	3.81	0.75	0.83	14.67**	0.24	2.75
	Innovation	*Fall 2017 Control*	*4.00*	*4.14*	*0.67*	*0.69*	*5.15**	*0.21*	–
		Spring 2018 Exptl	3.56	3.74	0.77	1.04	6.55*	0.20	0.21
		Fall 2018 Exptl	3.75	3.94	0.69	0.76	16.48**	0.26	1.23
		Spring 2019 Exptl	3.76	4.05	0.65	0.79	29.38**	0.55	4.82*
Attitudes toward science	Future intentions to study science	*Fall 2017 Control*	*3.07*	*3.15*	*1.16*	*1.17*	*1.51*	*0.07*	–
		Spring 2018 Exptl	2.62	2.84	1.15	1.15	9.00**	0.20	1.34
		Fall 2018 Exptl	2.96	3.11	1.14	1.13	7.29**	0.13	1.12
		Spring 2019 Exptl	2.77	3.06	1.02	1.06	20.34**	0.27	2.55
	Science self-efficacy	*Fall 2017 Control*	*3.82*	*3.87*	*0.73*	*0.79*	*1.08*	*0.06*	–
		Spring 2018 Exptl	3.55	3.66	0.77	0.87	3.80	0.13	1.97
		Fall 2018 Exptl	3.70	3.76	0.75	0.83	2.16	0.08	0.22
		Spring 2019 Exptl	3.66	3.72	0.69	0.77	1.54	0.11	1.31
	Scientific storytelling	*Fall 2017 Control*	*3.96*	*4.09*	*0.70*	*0.73*	*4.80**	*0.18*	–
		Spring 2018 Exptl	3.69	3.83	0.78	0.94	4.16*	0.16	0.27
		Fall 2018 Exptl	3.87	3.95	0.63	0.77	3.03	0.11	1.38
		Spring 2019 Exptl	3.94	4.07	0.67	0.81	6.10*	0.24	1.32
	Place-based learning	*Fall 2017 Control*	*3.70*	*3.84*	*0.70*	*0.81*	*4.49**	*0.19*	–
		Spring 2018 Exptl	3.53	3.63	0.76	0.88	2.16	0.12	1.74
		Fall 2018 Exptl	3.66	3.73	0.69	0.77	1.74	0.10	1.14
		Spring 2019 Exptl	3.72	3.86	0.67	0.75	6.66*	0.28	1.24

$n = 146$ (Fall 2017), 141 (Spring 2018), 204 (Fall 2018), 202 (Spring 2019) *$p < 0.05$, **$p < 0.0$.

Results

Quantitative data

For the control group (Fall 2017), pretest-posttest changes were fairly small, positive (except for *Personal Relevance*), and statistically significant for three scales (*Innovation*, *Scientific Storytelling*, and *Place-Based Learning*) at $p < 0.05$. Differences in QuACS scores between the pretest and posttest for the control group ranged from d = −0.03 standard deviations for *Personal Relevance* to 0.21 standard deviations for *Innovation*. For each of the three experimental groups, pretest-posttest changes were positive and relatively larger than for the control group for most QuACS scales. For the scales of *Personal Relevance* and *Future Intentions to Study Science*, pretest-posttest changes were statistically significant for each of the three experimental groups ($p < 0.01$). For the *Innovation* scale, pretest-posttest changes were significant for the Spring 2018 group ($p < 0.05$) and for the Fall 2018 and Spring 2019 groups ($p < 0.01$). For the *Scientific Storytelling* scale, a significant pretest-posttest difference ($p < 0.05$) emerged during Spring 2018 and Spring 2019, but not Fall 2018. Finally, a significant difference ($p < 0.05$) emerged for the *Place-Based Learning* scale for Spring 2019 only. As previously mentioned, calculations of these differences were accompanied by calculations of effect size to truly understand their significance. Across the three experimental groups, differences in QuACS scores between the pretest and posttest ranged from 0.08 standard deviations for *Science Self-Efficacy* to 0.55 standard deviations for *Innovation*. These values represent a range of small to medium effect sizes. Pretest-Posttest differences that were nonsignificant were associated with effect sizes ranging from only 0.08 standard deviations for *Science Self-Efficacy* (Fall 2018) to 0.13 standard deviations also for *Science Self-Efficacy* (Spring 2018), which represent negligible differences that are educationally unimportant.

A total of 18 ANCOVAs were used to compare the control group (Fall 2017) with each of the three experimental groups in terms of posttest scores on each of the six QuACS scales while controlling for pretest scores. The last column of Table 5 shows that ANCOVA results were nonsignificant for all of the 18 cases with the one exception of the *Innovation* scale for the Spring 2019 experimental group. That is, although significant pretest-posttest changes were found for numerous QuACS scales for each of the three experimental groups, the ANCOVA results generally suggest that pretest-posttest changes experienced by the experimental groups were not significantly larger than those experienced by the control group, with one exception. The one exception was that, according to ANCOVA, the Spring 2019 group's posttest *Innovation* scores were significantly superior to the control group's scores when pretest scores were controlled.

The faculty participants implementing SPAN assignments remained consistent across the semesters. Therefore, we anticipated that their levels of comfort, familiarity, and overall knowledge of implementation of SPAN assignments would increase over time. Given this, we hypothesized that the magnitude of pretest-posttest changes in QuACS scores would increase gradually over time from Fall 2017 to Spring 2019. The ANOVA results and effect sizes in Table 5 indicate that pretest-posttest changes for different QuACS scales for different groups:

- ranged from −0.03 to 0.21 standard deviations; significant for 3 scales, for the Fall 2017 control group
- ranged from 0.12 to 0.20 standard deviations; significant for 4 scales, for the Spring 2018 experimental group
- ranged from 0.08 to 0.26 standard deviations; significant for 3 scales, for the Fall 2018 experimental group
- ranged from 0.11 to 0.55 standard deviations; significant for 5 scales for the Spring 2019 experimental group.

The above results tentatively suggest that, although improvement over time was slow, the Spring 2019 group performed better than any earlier groups in terms of pretest-posttest changes on the QuACS (especially for the scale of *Innovation*).

Qualitative data results

Experiencing enjoyment

As interview data were analyzed, one of the primary themes to emerge was *enjoyment*. Across interviews, students consistently shared about their active engagement with the SPAN assignment, largely due to their finding the assignment enjoyable. One student said, "I put time and effort into it. I think I did a little more than others because it was a really fun assignment." Another student said, "When teachers make things fun, which I think is really difficult when you are in college, you want to invest time into it." Interview analyses also helped us determine the specific aspects of students' experiences with SPAN assignments that made them enjoyable. Among several interviews, students indicated that their ability to have significant control and choice in completing the assignment was of great importance and related to their overall enjoyment of the assignment.

For example, students frequently reflected on their ability to choose an area of focus, decide the style of audio narrative, and control the type of product they wished to create to fulfill the SPAN assignment expectations. This also relates to the ways that SPAN functions as a descriptive pedagogical innovation rather than the prescriptive implementation of an assignment within a course. One student said, "I think that is what students really want, the fact that you get to choose something [referring to topic and product] in a classroom and enjoy it and make the best out of it." Another student commented, "…because I chose the topic, I enjoyed every part of it." In addition to control and choice, most students expressed that their ability to engage with heightened forms of creativity was also key to their enjoyment. Students frequently commented on the importance of being able to express themselves, particularly in a science class, a category of classes that are usually highly structured and do not always welcome various forms of creativity. One student said, "It [SPAN] really let you be creative. I think that really promoted the students to want to invest their time into the assignment." Another student commented that, "You would be able to add music to it, for example, and your own kind of flavor to the project to make it your own." Beyond their own creativity, students also commented on the importance of being able to see how other forms of creativity emerged among their peers' assignments. As a result, students felt more connected to one another and more comfortable within the classroom space.

Experiencing increased levels of personal relevance

Analysis of interview data yielded a second prominent theme of *personal relevance*. This is aligned with the quantitative data, which revealed significant pretest-posttest differences with small effect sizes for each of the experimental groups of students. Throughout all the interviews, students reported that the SPAN assignment enabled them to make deeper connections between the assignment, the content, and their worlds outside of the classroom. One student described the experience saying, "I took a science class before and it was really boring and I really hated it. I didn't see the relevance at all with what I wanted to do with my education, so I kind of strayed away. After taking this class, I saw that it could be so much more fun with so much more to read and learn about." The theme of personal relevance is, naturally, also deeply personal. What one student might consider personally relevant might not be the same as the connections another student makes.

Our interview analyses helped to further uncover some of the dimensions of personal relevance that came to light as a result of students' engagement with SPAN assignments. First, several students discussed the unique connections made between the content learned as part of the curriculum and its importance and application in a real-world context. One student said, "I feel as though being able to do an audio [SPAN] assignment project, I enjoyed my science class more than if I just took a bunch of tests and quizzes. I actually learned something." Second, many students identified several transferrable skills that they developed and/or strengthened during the process of completing their SPAN assignment. For example, honing research skills, finding ways to communicate scientific information in compelling and creative ways, working with audio technology, engaging in a robust peer review process, and learning the fundamentals of conducting interviews were commonly discussed by students. Finally, students reported feeling exceptionally accomplished and proud of their progress and final SPAN product. This sense of pride was linked to the fact that students took the time to engage with their work, challenge themselves, make personal connections to science content, and learn new skills along the way.

SPAN as pedagogical innovation

Analyses of student interviews uncovered one final overarching theme, *innovation*. Similar to personal relevance, the theme of innovation is supported by positive changes in the quantitative data. Data from interviews suggest that students viewed SPAN assignments as a different type of learning tool than those utilized in a traditional science classroom. Students further described SPAN as a pedagogical innovation because it was unlike anything they had previously experienced or been asked to do in order to learn and connect with science content. One student emphatically stated, "Not in one science class have I ever had the ability to have as much creative opportunity." Another student put it simply; "There is always a bit of inspiration when it comes to doing new things." These "new things" are part of what students frequently discussed when describing SPAN as innovative. Specifically, learning new skills (as described earlier), breaking away from traditional forms of assessment typically associated with science courses, building an understanding of the nature of science, and furthering their understanding of different modalities for communicating scientific information, with a particular emphasis on the power of narrative.

Students' critiques of SPAN assignments

As part of our qualitative data collection, we also engaged students to hear their critiques and perceived limitations of SPAN assignments. Our analyses revealed two overarching critiques, namely issues with technology and, to a much lesser extent, content limitations. The first of these critiques was clustered primarily during the first semester of implementation of SPAN assignments. For example, students reported needing significant assistance in learning and working with the software and associated technology. This seemed to cause some initial frustration among students; however, students further reported persisting through this learning curve and ultimately enjoying the assignments.

Our analyses showed that critiques surrounding technology issues dwindled in frequency with each subsequent semester of implementation. This suggests that early in the study, faculty participants had not yet fully developed their own technical capacities and were still learning how to best teach those skills to their students. Further, the research team reacted to this early data and began to provide faculty with more targeted assistance related to troubleshooting the technology. As the study progressed, students' critiques shifted from technology toward content. (Separately, faculty reported an increased comfort in their teaching of the technical skills involved.) The scope of the critique was primarily rooted in struggling to ensure that content was being relayed accurately while maintaining an awareness of students' creative freedoms in communicating the content, as well as trying to remain focused on the topic under examination in order to meet assignment expectations, rather than trying to fit all the information found during the research process. Although students presented these latter ideas as critiques, we posit that these aspects of SPAN assignments were simultaneously responsible for students perceiving heightened levels of enjoyment, personal relevance, and innovation.

Discussion

The data presented in this paper support our hypothesis that SPAN assignments would have a positive impact on students' perceptions of the classroom learning environment and their attitudes toward science in introductory geoscience courses. Specifically, we saw consistent positive pretest-posttest changes in both learning environment scales, *Personal Relevance* and *Innovation*, and the attitude scale *Future Intentions to Study Science*. Effect sizes of these changes, indicating magnitude or educational importance, ranged from small to medium in size. Further, the magnitude of pretest-posttest changes in QuACS scores increased gradually over time from Fall 2017 to Spring 2019. Our data suggest that, although improvement over time was slow, the Spring 2019 group showed greater increases than the earlier groups in terms of pretest-posttest changes on the QuACS. We believe the increase in the QuACS scales showing significant pretest-posttest differences as well as an increasing trend toward larger effect sizes later in the study demonstrates that as the pilot faculty improved their implementation, the SPAN assignments became more understandable, useful, and powerful. This resulted in increases in aspects of the students' perceptions of the classroom learning environment and attitudes toward science.

Leading up to the final semester of implementation, the SPAN assignments did not appear to have a significant influence on the *Place-based Learning, Scientific Storytelling, and Science Self-Efficacy* scales; this was to be expected. For example, SPAN assignments were not specifically designed to be used in place-based learning. Some instructors used placed-based assignments (e.g., "explore your local watershed") while others did not. Within this study, we did not attempt to determine how students defined "their place" or "their community." Regardless, some growth in this area may be expected as students found the material more personally relevant, and so there could be spillover into a sense of place or community. In the area of learning science through storytelling, the control and two of the implementation semesters are very similar, although the final semester of implementation did have a larger effect size. Scientific storytelling can be a component of any science course; instructors often tell stories in lecture and incorporate local examples, possibly causing similarities in how students interpreted this scale during the control semester and subsequent implementation semesters. In addition, SPAN assignments did not appear to significantly increase students' sense of self-efficacy, or the extent to which students believe that they can be successful in science. This may be because students viewed the assignments as novel and "unexpected" in a science course, and so perhaps the connections and communication skills they developed seemed unrelated to the "work" of a scientist. This study shows that SPAN assignments, on their own, did not enhance these three areas.

Limitations

In pilot studies such as this the variability designed to explore the parameter space also creates some challenges in analysis. For example in studies like this, control populations are extremely difficult to establish. We attempted to use the Fall 2017 semester to serve as a control. However, this approach had several challenges that make the comparison to other semesters difficult, and therefore impacted the relevance of the ANCOVA *F* values. In Fall 2017, the QuACS survey evaluated the change in student perceptions at the beginning and end of courses that did not contain a SPAN assignment. Pilot faculty attempted to make these courses as similar as possible to those in which they would be implementing SPAN, but matching student population, specific courses, and experiences is nearly impossible. A more accurate control would be to have a large class, with the same instructor, course material, and student population, in which half of the students complete SPAN and half complete another type of assignment. This was not feasible given the objective to pilot the SPAN approach within a variety of environments, including smaller class sizes.

Variability between semesters is not just a factor of instructors increasing their confidence in implementing SPAN. Instructors often taught different courses in different semesters (for example, introduction to geology in the fall and oceanography in the spring). In this pilot study, we aimed to capture a diverse range of courses and implementation approaches. Our study supports widespread effectiveness in SPAN assignments, helping students increase their senses of personal relevance, curricular innovation, and intentions to study science in the future, but the consequence of these diverse environments is the corresponding increase in the number of variables. Therefore, we suggest that looking at the final two semesters together is a good average of the impact. In Spring 2018, the first implementation semester, faculty interviews and data show that pilot faculty were still working to establish SPAN into their courses, so the students likely received only partially developed assignments. By Fall 2018 and Spring 2019, the faculty felt confident and made only minor adjustments to their assignments (based on expected variables such as changes in schedule, slightly different course content, variations in class size, etc.). Therefore, in the final two semesters, we believe that students received "fully developed" SPAN assignments, and the variability between the semesters captures the typical variability of a full academic year (variations in the types of courses taught and students enrolling in spring versus fall).

An additional limitation of this research is related to the low sample size of the semi-structured interviews. As noted earlier, we successfully conducted 9 interviews from an available sample of 54 students. Despite the small sample size, we assert that the data derived from the interviews is trustworthy, as it also aligns with and is supported by our results from the quantitative data analyses. For future research, we will strive to increase the percentage of semi-structured interviews completed.

Finally, a limitation is that these assignments occur within the context of a full course. Students are having other experiences with the content and instructor. SPAN assignments, by design, are not curricular changes to courses. And courses, themselves, occur within the context of other educational experiences. Each course itself has some level of variability within a semester with regard to SPAN implementation. The nature of a SPAN assignment allowed each instructor control over their implementation of the SPAN framework. Thus, there was some variability in implementation. For example, some instructors devoted more or less time on the assignment. Others provided more or less scaffolding for students. This variability may have some impact on the outcomes and subsequent effect sizes. The addition of qualitative interviews helps mitigate some of these limitations. In addition, the use of the same group of pilot faculty/institutions and course types helps control for some variability. We have proposed future work to address some of these limitations and better differentiate the impact of SPAN assignments.

Conclusions

Given the documented need for increased representation in the geosciences and the opportunities that introductory courses provide, it is important to make deliberate changes to introductory teaching with an awareness of the specific challenges. This requires a pedagogical shift in teaching, one that is flexible, in order to address the challenges, and inclusive, in order to take advantage of the opportunities. We cannot reach new, larger, and more diverse student bodies with the same approaches. Collectively we must engage in pedagogical shifts to make deliberate change in teaching pedagogy. Taken in whole, SPAN assignments show great promise as a pedagogical tool that is low-barrier-to-entry and realistic for improving introductory students' connections to geoscience content. Even today, many introductory-level courses use a lecture approach with a focus on presenting content and evaluating students' ability to reorganize and replicate that content through tests and exams (Egger, 2019b). This pilot study documents that SPAN assignments can be successfully implemented in a wide variety of introductory level geoscience courses. They increase positive student perceptions of their learning environment in the areas of personal relevance, curricular innovation, and future intentions to study science. Interviews indicate that this was related to the choice and creativity embedded in the design of a SPAN assignment. These findings stretched across a range of course types and institutions, indicating promise for wide-spread implementation in introductory-level geoscience courses.

Acknowledgments

We extend deep gratitude to our pilot faculty and institutions for piloting this program. We acknowledge Dr. Barry Fraser for assistance with and useful discussions about quantitative statistical analysis. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Funding

This work was supported by the National Science Foundation under Grant #1708590. G. Simmens received undergraduate researcher support from Kutztown University. IRB approval for human subject research was granted through Kutztown University #02092016.

References

Anderson, G. (1998). *Fundamentals of educational research* (2nd ed.). Falmer Press.

Boslaugh, S., & Watters, P. A. (2008). *Statistics in a nutshell: A desktop quick reference*. O'Reilly Media.

Campbell, G. (2005). Podcasting in education. *EDUCAUSE Review*, *40*, 33–46.

Clough, M. P. (2011). The story behind the science: Bringing science and scientists to life in post-secondary science education. *Science & Education*, *20*(7-8), 701–717. https://doi.org/10.1007/s11191-010-9310-7

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Cortina, J. (1993). What is coefficient alpha?: An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98–104. https://doi.org/10.1037/0021-9010.78.1.98

Dahlstrom, M. F. (2014). Using narratives to communicate science with nonexpert audiences. *Proceedings of the National Academy of Sciences USA*, *111*(4), 13614–13620. https://doi.org/10.1073/pnas.1320645111

Deal, A. (2007). *Teaching with technology: Podcasting*. Eberly Center, Carnegie Mellon University. https://www.cmu.edu/teaching/technology/whitepapers/Podcasting_Jun07.pdf

Dhingra, K. (2008). Towards science educational spaces as dynamic and coauthored communities of practice. *Cultural Studies of Science Education*, *3*(1), 123–144. https://doi.org/10.1007/s11422-007-9077-6

Dorman, J. P., & Adams, J. E. (2004). Associations between students' perceptions of classroom environment and academic efficacy in Australian and British secondary schools. *Westminster Studies in Education*, *27*(1), 69–85.

Drew, C. (2017). Edutaining audio: An exploration of education podcast design possibilities. *Educational Media International*, *54*(1), 48–62. https://doi.org/10.1080/09523987.2017.1324360

Egger, A. (2019a). The role of introductory geoscience courses in preparing teachers—and all students— for the future: Are we making the grade?*GSA Today*, *29*(10), 4–10. https://doi.org/10.1130/GSATG393A.1

Egger, A. E. (2019b). Introductory geoscience courses and STEM education research. *Journal of Geoscience Education*, *67*(3), 193–194. https://doi.org/10.1080/10899995.2019.1633876

Emdin, C. (2010). *Urban science education for the hip-hop generation*. Sense Publishers.

Epstein, A., Easton, J., Murthy, R., Davidson, E., de Bruijn, J., Hayse, T., Hens, E., & Lloyd, M. (2010). *Helping engineering and science students find their voice: Radio production as a way to enhance students' communication skills and their competence at placing engineering and science in a broader societal context*. Proceedings of the American Society for Engineering Education Annual Conference (pp. 15.640.1–15.640.13).

Fraser, B. J., & Tobin, K. (1991). Combining qualitative and quantitative methods in classroom environment research. In B. J. Fraser & H. J. Wallberg (Eds.), *Educational environments: Evaluation, antecedents and consequences* (pp. 271–292). Pergamon.

Gardiner, L. (2018, December 10–14). *Science in anNutshell: Engaging the public with geoscience using creative nonfiction storytelling*. Presented at the 2018 Fall Meeting [PA22C-05], AGU, Washington, DC.

Guertin, L. (2012). Community-based research projects with podcasting in introductory-level geoscience courses. *Abstracts with Programs (Geological Society of America)*, *44*(7), 611.

Holm Adamec, B., Asher, P. (2013). *The important role of two-year colleges in the earth and space sciences*. American Geophysical Union. https://www.agu.org/-/media/Files/Learn-and-Develop/AGU_URECAS_Workshop_Report_2013.pdf?la=en&hash=B6E1E991E83B0483396E41B25D3AAFCBAB492D0D

Keat, R., & Urry, J. (1982). *Social theory as science* (2nd ed.). Routledge & Kegan Paul.

Klassen, S. (2009). The construction and analysis of a science story: A proposed methodology. *Science & Education*, *18*(3-4), 401–423. https://doi.org/10.1007/s11191-008-9141-y

Kleege, G., & Wallin, S. (2015). Audio description as a pedagogical tool. *Disability Studies Quarterly*, *35*(2). https://doi.org/10.18061/dsq.v35i2.4622

Kraal, E. R., & Regensburger, M. (2013). Telescopic topics: Student-created radio podcasts in a general education planetary science class. *Abstracts with Programs (Geological Society of America)*, *45*(7), 469.

Lee, M. J. W., & Chan, A. (2007). Reducing the effects of isolation and promoting inclusivity for distance learners through podcasting Turkish online. *Turkish Online Journal of Distance Education*, *8*(1), 85–105.

Lee, M. J. W., McLoughlin, C., & Chan, A. (2008). Talk the talk: Learner-generated podcasts as catalysts for knowledge creation. *British Journal of Educational Technology*, *39*(3), 501–521. https://doi.org/10.1111/j.1467-8535.2007.00746.x

Lundby, K. (2008). *Digital storytelling, mediatized stories: Self-representations in new media*. Peter Lang.

Martinez, C., & Baker, M. A. (2006). *Introductory geoscience enrollment in the United States: Academic year 2004-2005 (GW-06-001)*. American Geosciences Institute.

Morgan, D. L. (1997). *Practical strategies for combining qualitative and quantitative methods*. Portland State University Press.

National Academies of Sciences, Engineering, and Medicine. (2015). *Integrating discovery-based research into the undergraduate curriculum: Report of a convocation*. The National Academies Press. https://doi.org/10.17226/21851

National Academies of Sciences, Engineering, and Medicine. (2017). *Service-learning in undergraduate geosciences: Proceedings of a Workshop*. The National Academies Press. https://doi.org/10.17226/24621

Nuñez, A.-M., Rivera, J., & Hallmark, T. (2020). Applying an intersectionality lens to expand equity in the geosciences. *Journal of Geoscience Education*, *68*(2), 97–114. https://doi.org/10.1080/10899995.2019.1675131

O'Neill, T., & Calabrese-Barton, A. (2005). Uncovering student ownership in science learning: The making of a student created mini-documentary. *School Science and Mathematics*, *105*, 292–301.

Pegrum, M., Bartle, E., & Longnecker, N. (2015). Can creative podcasting promote deep learning? The use of podcasting for learning content in an undergraduate science unit. *British Journal of Educational Technology*, *46*(1), 142–152. https://doi.org/10.1111/bjet.12133

Pew Research Center. (2019). *Audio and podcasting fact sheet*. https://www.journalism.org/fact-sheet/audio-and-podcasting/

Robin, B. R. (2006). The educational uses of digital storytelling. *Society for Information Technology and Teacher Education International Conference*, *2006*(1), 709–716.

Robin, B. R. (2008). Digital storytelling: A powerful technology tool for the 21st century classroom. *Theory into Practice*, *47*(3), 220–228. https://doi.org/10.1080/00405840802153916

Rubin, H. J., & Rubin, I. S. (2005). *Qualitative interviewing: The art of hearing the data* (2nd ed.). SAGE Publications.

Schmidt, R. (2013). *Bridging the geoscientist workforce gap: Advanced High School Geoscience Programs* [Ed. D. thesis]. Drexel University.

Sexton, J. M., Pugh, K. J., Bergstrom, C. M., & Riggs, E. M. (2018). Reasons undergraduate students majored in geology across six universities: The importance of gender and department. *Journal of Geoscience Education*, *66*(4), 319–336. https://doi.org/10.1080/10899995.2018.1507546

Sirrakos, G., & Fraser, B. (2019). Development and validation of the questionnaire assessing connections to science (QuACS). In D. B. Zandvliet and B. J. Fisher (Eds.), *Thirty years of learning environments: Looking back and looking forward (Advances in learning environments research series)* (pp. 102–126). Brill Sense. https://doi.org/10.1163/9789004387720_006

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson.

Tobin, K., & Fraser, B. J. (1998). Qualitative and quantitative landscapes of classroom learning environments. In B. J. Fraser, K. G. Tobin, & C. J. McRobbie (Eds.), *Second international handbook of science education* (pp. 623–640). Springer.

Wilson, C. (2014). *Status of the geoscience workforce*. American Geosciences Institute.

Wilson, C. (2016). *Status of the geoscience workforce*. American Geociences Institute.