Received: 14 October 2021 | Revised: 27 December 2021 | Accepted: 6 January 2022

DOI: 10.1111/ssm.12515

RESEARCH PAPER - SCIENCE EDUCATION

WILEY | SSMA

Design of a phenomenon-based science outreach program and its effects on elementary students' epistemological understanding of, and attitudes toward, science

Ryan T. Helsel[1] | **Sammi Lambert**[1] | **Lindsey Dickerson**[2] | **Jack Strelich**[2] | **Vanessa Woods**[2] | **Darby Feldwinn**[1,3]

[1]Department of Chemistry and Biochemistry, University of California, Santa Barbara, California, USA

[2]Department of Psychological and Brain Sciences, University of California, Santa Barbara, California, USA

[3]Department of Education, University of California, Santa Barbara, California, USA

Correspondence
Darby Feldwinn, Department of Chemistry and Biochemistry, University of California Santa Barbara, Santa Barbara, CA 93106-9510, USA.
Email: feldwinn@chem.ucsb.edu

Funding information
California Subject Matter Project; Every Student Succeeds Act; Karisma Foundation

Abstract

This study describes the design and implementation of a science outreach program (elementary; 2nd–5th) and its associated student outcomes. Key features of the program include: (a) phenomenon-based, NGSS-aligned curriculum designed by science education experts and influenced by educational, sociocultural, and cognitive science theories, (b) active engagement by elementary students in experimentation and exploration of a common phenomenon (2 per year, called modules) across multiple sessions (6–8 per module), (c) professional development (PD) provided to scientists (mentors) and teachers to supply classrooms with multiple (5–7) adult more knowledgeable others (AMKOs), and (d) close classroom interaction between elementary students and mentors, allowing for students to participate in evidence-based sense-making through whole-class and small-group discussions. We examined the effects of program participation on students' epistemological understanding of science (EUS) and attitudes toward science, as well as surveyed teachers to determine how they felt the program affected students' attitudes and EUS. Student measures indicated they developed a deeper understanding of the process of scientific knowledge generation and were more likely to report liking science more than less. Similarly, teachers' self-reports corroborated these results, as well as showed teachers thought the structure of the program effectively integrated mathematics and language arts.

KEYWORDS

attitudes, cooperative learning, curriculum development, inquiry, science education

1 | INTRODUCTION

Science, Technology, Engineering, and Math (STEM) interest is commonly first reported in elementary school (Dabney et al., 2013; Maltese & Tai, 2009), with students' aspirations to pursue STEM careers solidifying as early as middle school (Sadler et al., 2012). Students typically lose interest in STEM with age, making the elementary years a target for early intervention programs that spark and maintain interest in science (Krapp & Prenzel, 2011).

To improve science education, California adopted the Next Generation Science Standards (NGSS) in 2013, which embraced a more holistic way of teaching science that embeds content within Science and Engineering Practices (SEPs) and Crosscutting Concepts (CCCs; NGSS Lead States, 2013). However, the California Science Test

(CAST) results revealed less than a third of 5th grade students met the standards (California Department of Education, 2019), indicating a need for more effective implementation of NGSS in elementary science education.

In addition to traditional classroom curriculum/teaching, outreach programs can play a role in designing and presenting NGSS-aligned curriculum in the classroom to foster a deeper understanding of, and interest in, science. Science outreach programs can be categorized into two types: Scientist in the Classroom Outreach (SCO) and Inquiry Outreach (IO) programs. SCO programs feature scientists sharing personal experiences and demonstrating/doing hands-on, non-three-dimensional (3D) activities (e.g., Clarke et al., 2019). IO programs have scientists design and train teachers to use curriculum (not necessarily aligned to standards) in their classroom, generally without program support (e.g., Cotabish et al., 2013). Both types of programs improve students' attitudes toward science (e.g., Clarke et al., 2019; Patrick et al., 2008). In addition, IO programs improve students' content knowledge (typically not assessed in SCO; Cotabish et al., 2013; Patrick et al., 2008). However, neither program assesses epistemological understanding of science (EUS). SciTrek is an outreach program that blends the strengths of SCO and IO programs: curriculum (modules) designed by science education specialists (IO strength), training for teachers (IO strength), and multiple mentors guiding students during modules that occur within school instructional time (SCO strength).

1.1 | Theoretical framework

1.1.1 | 3D teaching and inquiry theories

The SciTrek curriculum and program structure was informed by educational, sociocultural, and cognitive science theories. There are two common educational models: the older inquiry model, and the newer 3D model, which is aligned with the NGSS. The inquiry model is a method of engaging students in an authentic scientific process and received traction in the 1990's (e.g., American Association for the Advancement of Science, 1993). Pedaste et al. (2015) summarized this process in terms of inquiry stages which consisted of: orientation, conceptualization, investigation, conclusion, and discussion. Despite its popularity, doubts remain about the effectiveness of inquiry-based science education, with criticisms of the peripheral role of the teacher (Hmelo-Silver et al., 2007) and the necessity for inquiry practice to be more formally linked with content (Krajcik et al., 2014). Such criticisms are rectified in the 3D model (National Research Council, 2012; NGSS Lead States, 2013). Within the 3D model, students simultaneously engage in Disciplinary Core Ideas (DCIs; science content), SEPs, and CCCs. This combination is necessary for building an integrated network of knowledge and skills across multiple domains of science (e.g., Krajcik et al., 2014).

The SEPs in NGSS (the singular aspect of the original inquiry stages) are composed of eight practices that are critical for students to understand and implement to know how scientific knowledge is created (NRC, 2012). This SEP framework and inquiry stages were integral to the design of the curriculum utilized by SciTrek. During modules, students engage in all SEPs to learn science content while taking part in authentic science experiences, which is consistent with 3D teaching. In addition, concentrated instruction is provided on one practice per grade.

1.1.2 | Sociocultural theories

Sociocultural theory states that learning occurs in a sociocultural context (e.g., Mortimer & Scott, 2003), which is necessary to create an authentic scientific experience. This requires identifying the context in which learning is occurring and allowing students to engage in the material in different social environments (Cobb & Yackel, 1996; Tudge & Scrimsher, 2003). To accomplish this, a program must provide opportunities for students to discuss content in small groups/whole class with their peers, mentors, and teacher. While all these people can be more knowledgeable others (MKOs; Rogoff, 1995; Tudge & Scrimsher, 2003), adult MKOs (AMKOs) are essential because they can facilitate meaning-making for students via scaffolded discussions, which is an integral aspect of 3D teaching (Bybee, 2011). The presence of multiple AMKOs allows for simultaneous higher-level discussions with students to take place, which is not possible with a single teacher (adult) in the classroom.

SciTrek's program structure and curriculum embrace a social learning environment and provide the opportunity for scientific knowledge generation via active engagement in evidence-based small group, and whole class, discussions, which is a more representative reflection of the practices and norms of real-world science. Because mentors and teachers (AMKOs) are a crucial aspect in fostering a successful sociocultural environment, they are provided with professional development (PD) to help them effectively facilitate classroom discussions.

1.1.3 | Cognitive theories

In addition to sociocultural theory, cognitive science models informed SciTrek's curriculum design. Furtak et al. (2012) defined four cognitive domains: conceptual (C; knowledge base), procedural (P; methodological steps to build scientific knowledge), epistemic (E; understanding how scientific knowledge is created and is malleable), and social (S; how communication and collaboration foster the construction of scientific knowledge). In their meta-analysis, they found that lessons that utilized only the C domain (i.e., traditional, content-focused instruction) produced smaller effect sizes compared to lessons that include CPES domains; further, the large effect size of CPES-focused lessons was attributed to students participating in the process of generating, developing, and justifying explanations (Furtak et al., 2012).

The SciTrek modules are designed to ensure students engage in all domains (CPES) multiple times: basing modules on an NGSS-aligned common phenomenon (C); allowing small groups to design their own experiment (P); having groups contribute to the understanding of a common phenomenon through collaboration and engagement in argument from evidence (E); and allowing students to engage in AMKO-scaffolded, and peer-to-peer, discussions (S).

Combining the strengths of IO and SCO programs while incorporating guiding theories has resulted in distinct features within the SciTrek program: (a) phenomenon-based, NGSS-aligned curriculum designed by science education experts and influenced by educational, sociocultural, and cognitive science theories, (b) active engagement by students in experimentation and exploration of a common phenomenon (2 per year) across multiple sessions (6–8 per phenomenon, called modules), (c) PD for scientists (mentors) and teachers to provide classrooms with multiple (5–7) AMKOs, and (d) close interaction between students and mentors. Our research questions are: How does program participation affect (a) students' changes in liking of science (attitudes) and EUS? (b) teachers' perceptions of students' experiences?

1.2 | SciTrek program overview

SciTrek started in 2010 by two professors in the Department of Chemistry and Biochemistry at a research institute on the Pacific Coast of the United States, scitrek.chem.ucsb.edu. The elementary portion of the program works with 2nd–5th grades. During a module, mentors (89% undergraduates [88% of whom are STEM majors], 3% graduates, and 8% other) guide students through the process. Before AMKOs work in the classroom, they undergo an immersive, module-specific 1.5-h PD which allows them to engage in the module as learners and develop questioning and scaffolding techniques. Mentors facilitate the same group(s) during the module, helping form deeper bonds with students as they investigate a phenomenon. Group sizes and number of mentors vary depending upon grade (Figure 1).

SciTrek requires active involvement from classroom teachers, which includes taking part in PD (both in and out of their classroom), via an apprenticeship model, to develop their 3D teaching skills. Typically, teachers assume the role of the SciTrek lead (leads class discussions and oversees mentors) after year 3. During year 1, teachers (novices) complete the same PD as the mentors. During years 2–3, teachers' out-of-classroom PD consists of engaging in communities of practice (up to 5 teachers) to develop their confidence, and knowledge, with SciTrek's guidance. These skills are applied, in their classroom (with their students), with SciTrek staff providing coaching/PD. Once teachers have become experts (lead modules on their own, ~year 4), they provide mentorship during SciTrek's communities of practice PD.

1.3 | SciTrek program structure and curriculum

The program consists of one mini module, and two full-length modules (referred to as "modules") per class during the academic year. Mini modules are one-hour and occur in fall. Modules occur over multiple (6–8) one-hour sessions, two to three times per week during standard instructional time. Mini modules are completed as a whole class and facilitated solely by the lead, whereas modules are in small groups. To ensure a productive social learning environment and authentic science experience, SciTrek includes both whole class discussions (facilitated by the lead) as well as small group work (facilitated by mentors). In 2nd/3rd grades, there is one mentor per group, whereas in 4th/5th grades, mentors rotate between multiple (2–3) groups. Figure 1 (left panel) outlines module schedule, timing, and personnel.

1.3.1 | Mini module and module design

Mini modules and modules of a given grade level are designed to interconnect to help students learn content and deeply explore one specific SEP-like practice. Mini modules are the first introduction students have to the SEP-like practice that is highlighted for their grade (Figure 1, right panel). During modules, students engage in all SEPs, but receive explicit instruction on their grade level highlighted SEP-like practice. Modules are designed around an anchoring phenomenon for which there can be multiple manipulatable experiments. This means there are multiple variables the students can choose to change, ensuring groups will be planning and carrying out unique

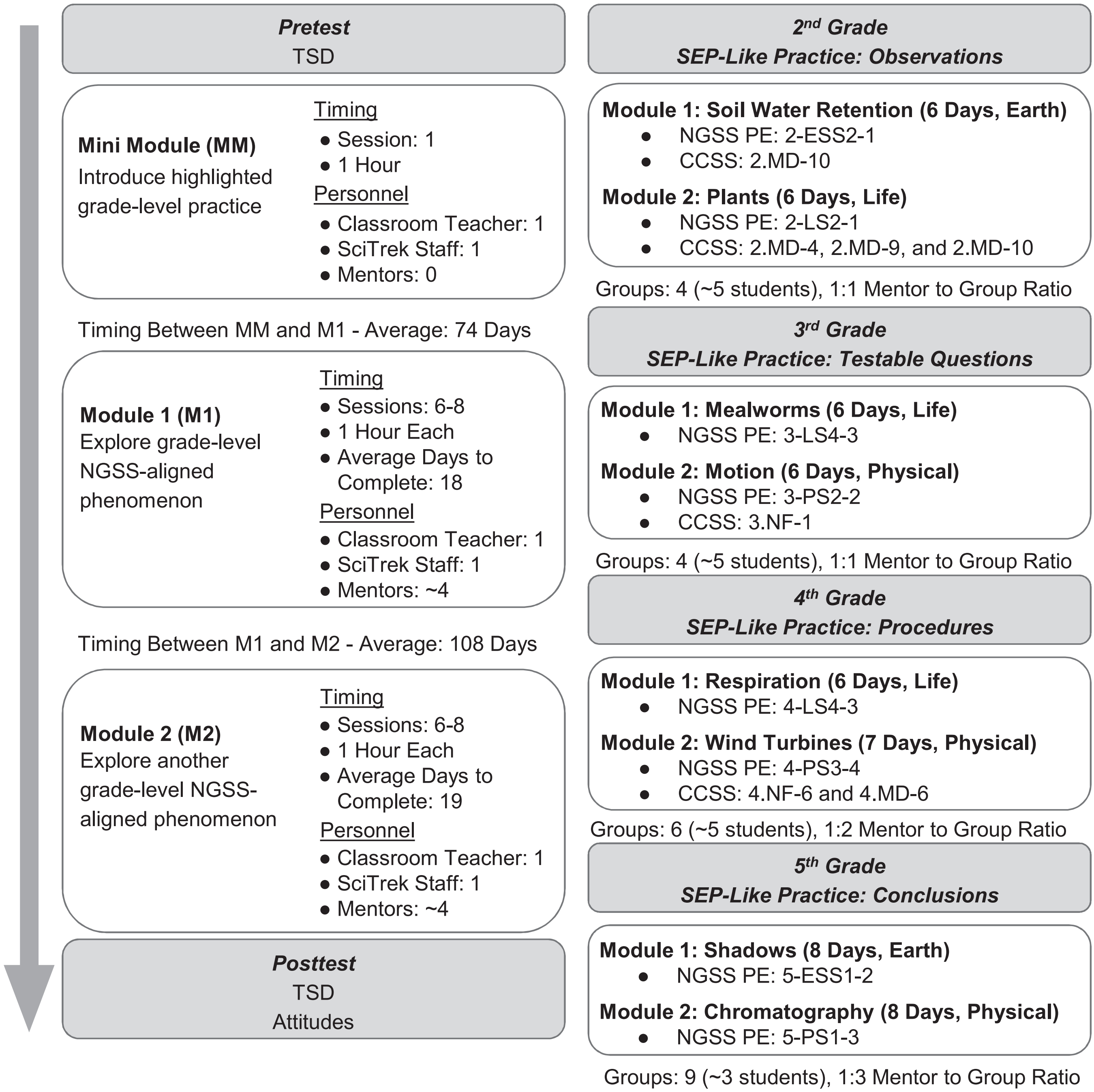


FIGURE 1 SciTrek module timing. Left column: Pictorial representation of a typical SciTrek schedule during an academic year including when particular measures were being assessed: Attitudes and Things Scientists Do (TSD)/Epistomological Understanding of Science (EUS) measure. Right column: Modules offered by SciTrek by grade with related Next Generation Science Standards (NGSS) Performance Expectation (PE) and Common Core State Standards (CCSS; when applicable). Mentor to group ratios are the minimum number of mentors that are present when doing modules. In addition, a SciTrek lead is also present in the classroom

experiments. This essential feature of SciTrek allows for all students to serve as MKOs, contributing to a deeper understanding of the phenomenon. Each phenomenon is chosen to highlight a specific NGSS Performance Expectation (PE) and, when applicable, integrate mathematics Common Core State Standards (CCSS; Figure 1). All SciTrek modules have integrated grade level appropriate writing and mathematics practices, even if a specific CCSS mathematics or language arts standard cannot be directly tied to the lesson sequence.

1.3.2 | Mini modules

The mini modules provide students with an introductory activity to their grade level highlighted SEP-like practice, which includes contextualizing, and implementing, scientific vocabulary through practice. Most are modified activities, which explore a specific scientific practice (e.g., 2nd grade; observations; potato candle modified from McIntosh, 2003).

1.3.3 | Full-length modules

Activities during modules link to inquiry stages, 3D learning (SEPs), sociocultural theory, and cognitive domains, and all modules have a similar progression of activities (Table 1). Students are given module-specific notebooks to keep detailed records. In 2nd and 3rd grades, to reduce the cognitive load and increase the engagement of Multilingual Learners (MLLs), mentors use large format notepads to write/draw students' thoughts and help students adjust ideas and transfer information to their notebooks.

Observe phenomenon

Each module begins with students in groups, led by a mentor, observing a common phenomenon (e.g., 2nd grade Soil Water Retention [SWR] module has students observe how soil compactness affects the amount of water absorbed), often tying to a real-world phenomenon (e.g., landslides). Students begin by making observations of the materials (Figure 2a), including discussing the scientific tools they will use, and having the mentors introduce contextualized, module-specific vocabulary (e.g., absorb, compact). As students observe the common phenomenon, mentors help them perform relevant calculations (e.g., amount of water absorbed by soil). After, the whole class is coached to relate these observations back to the real-world applications of the phenomenon, with the goal being the guided generation of a class research question (e.g., What variables affect how much liquid a soil can absorb?)

Generate variables

In 3rd–5th grades, mentors next help students brainstorm variables that could be manipulated, as well as facilitate discussions about how these manipulations may affect the phenomenon. Students listen and respond to their classmates' predictions. This allows students to collaborate, practice using new scientific vocabulary, and deepen their EUS. Providing scaffolding to students to help them predict how these variables affect the phenomenon assists them in selecting appropriate controls for their novel experiment.

Design experiment

Student groups then generate their specific research questions by selecting a changing variable for their experiment. In 2nd and 3rd grade, students are presented with ~3 variables to choose from. In 4th and 5th grade, students select any variable presented to them on the materials page, which is purposefully curated to give students a variety of materials while not inhibiting their experimental freedom (Chromatography [using paper and a liquid to smear out the colors in ink] materials page, Figure 2b). For SWR, a group research question could be, "If we change the soil type, what will happen to the amount of liquid that the soil absorbs?" Mentors then guide groups to determine their experimental set-up, including discussion and creation of a list of their controls and changing variable values. Mentors scaffold this process by asking questions (e.g., Why did you pick the changing variable values that you did?). These guiding questions and resultant revisions help groups think deeply about their experiment before it begins, which engages students' social and epistemic domains.

After forming their experimental set-up, groups are mentored in writing a procedure. In 2nd grade, procedures are drawn as a series of pictures with annotations, including control and changing variable values, as well as the data that will be collected (Figure 2c). Starting in 3rd grade, students generate written procedures. The mentor guides students through the process by asking questions (e.g., What steps do you remember from the phenomenon experiment?). Students then set up a results table and engage in scaffolded discussions to make predictions about what will happen during their experiment.

Conduct experiment

Students conduct the experiment by following their procedure and recording their measurements/observations in the results table. While each trial is usually individually assigned, all students collaborate in making observations and comparing between trials. After collecting data, students are differentially scaffolded across the grade levels to make a bar graph, which aids in analyzing their data for trends. In 2nd/3rd grades, students graph the results of their trial on an individual graph column (Figure 2d, left) to make it easier for the group to arrange the columns in increasing order. These column pieces are taped together to create a group graph (Figure 2d, right). In 4th/5th grades, the mentors guide their groups through a series of steps to directly create their bar graphs in increasing order. This process helps students realize that not all ways of representing data are equally useful for noticing trends, and gives them insight into the process of creating scientific knowledge. Graphing the data in ways that highlight trends makes it easier for students to draw a conclusion that is informed by data and allows them to serve as MKOs that can convince their peers of the validity of their findings. This mentored, social learning process builds their epistemic knowledge.

TABLE 1 Progression of activities during SciTrek modules for all grade levels

Module activity	Description	SEPs students engage in	Stage[a] (CPES)[b]
Observe phenomenon	Groups make observations of a common module-specific phenomenon using appropriate scientific tools	1. Asking questions (for science) and defining problems (for engineering)	1 (PS)
Generate variables	Groups brainstorm ways they can manipulate the phenomenon to answer the class question	1. Asking questions (for science) and defining problems (for engineering)	2 (S)
Design experiment	Groups decide on an aspect of the phenomenon to explore, create a research question, and design an experiment. While not required, typically groups explore different aspects of the phenomenon resulting in different experiments	3. Planning and carrying out investigations	3a (PES)
Conduct experiment	Groups conduct their experiment, collect data, and graph their results	3. Planning and carrying out investigations 5. Using mathematics and computational thinking	3b (PES)
Revise experiment	Students analyze mock data to determine why having more than one changing variable is problematic. Based on this knowledge, they repeat stages 3a and 3b. (5th Grade Only)	3. Planning and carrying out investigations 5. Using mathematics and computational thinking	3c (PES)
Form conclusion	Groups analyze their results to make a claim and use supporting data to aid in explaining an aspect of the phenomenon	4. Analyzing and interpreting data 6. Constructing explanations (for science) and designing solutions (for engineering)	4a (CPES)
Create and present posters	Groups create a poster to share their inquiry cycle. Students are encouraged to ask questions of the presenting group. In addition, AMKOs ask guiding questions. These questions help the class analyze their data, make connections between presentations, and gain a deeper understanding of the phenomenon	7. Engaging in argument from evidence 8. Obtaining, evaluating, and communicating information	4b (CES)
Connect experiments to the NGSS	With guidance from the lead, students' results are connected back to the NGSS by applying other scientists' data to develop a more comprehensive model of the phenomenon	2. Developing and using models 4. Analyzing and interpreting data	4c (CE)

Notes: Students work with the same group (3–5 students) for the entire module and are completing these activities in a mentored social context, which is consistent with sociocultural theory.

[a]Inquiry stage 1 (orientation), 2 (conceptualization), 3 (investigation), and 4 (conclusion).

[b]C (conceptual), P (procedural), E (epistemic), and S (social).

Revise experiment (5th grade only)

In 5th grade, the highlighted SEP-like practice is conclusions. When designing the first experiment, 5th grade students are not told the number of changing variables to use, and most groups choose to change multiple variables for their first experiment. After groups complete their experiments, the class analyzes mock data to arrive at the consensus that if there are multiple changing variables, they cannot make a conclusion. This allows students to increase their epistemological knowledge and learn about conclusions through authentic science practice. Students then analyze their data and make a conclusion or state why a conclusion cannot be made. After the discussion, groups undergo the meaningful process of revision by designing and running a new experiment with a single changing variable. The revision process allows students to understand that mistakes in experiments can be valuable.

Form conclusion

Groups then analyze their data using their graph to make a claim about how their changing variable affects their measurements. Then, they support that claim with evidence (i.e., experimental data). The mentor guides the group to select the strongest evidence to support their claim, and the group's conclusion is the product of collaboration and discussion.

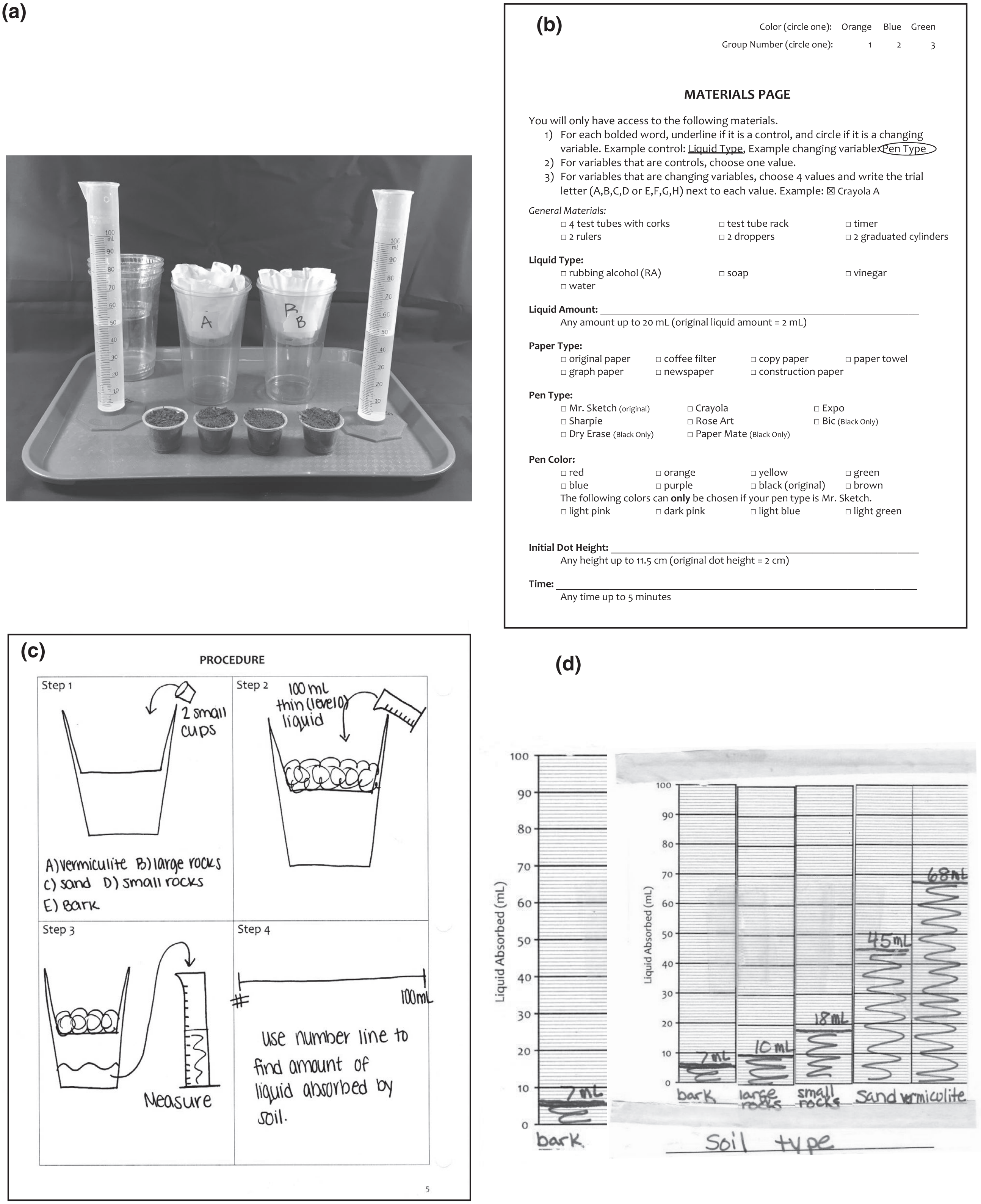


FIGURE 2 SciTrek module examples. (a) Soil water retention (SWR) experimental set-up, (b) chromatography materials page, (c) SWR procedure, (d) SWR individual graph piece (left) and five individual graph pieces taped together to make full graph (right). For more details on modules, see program website, scitrek.chem.ucsb.edu

Students are then asked to discuss ways in which they acted like scientists, besides carrying out the experiment. Mentors guide students in generating unique, specific ideas that highlight practices they completed (e.g., asking questions, writing procedures, and collaborating with their group). This allows students to reflect on how scientific knowledge is generated.

Create and present posters

Each group makes a poster of their investigative process by displaying their research question, describing how their experiments were structured, and presenting their data and conclusion. The posters are presented to the class in order of increasing complexity, with posters about the same changing variable presented back-to-back. Notes are taken by the lead (all grades) and students ($4^{th}/5^{th}$ only), which allows students to compare results across multiple groups. During these presentations, students are encouraged to participate in a feedback cycle regarding the experiments to analyze each group's results by asking questions of the presenting group. If students do not make connections between the experiments on their own, AMKOs ask the group guiding questions to connect the presentation both to other presentations, and to the common phenomenon. Each presentation culminates in the class generating a summary of what they learned. The questioning during the presentations fosters reflection and allows students to engage in self-evaluation to identify which aspects of the phenomenon they do not yet fully understand, thereby developing their social and epistemic domains.

Connect experiments to the NGSS

During the final session, the experiments are tied back to the NGSS PE for the module, and students are provided scaffolding in developing a more comprehensive cognitive model of the phenomenon via whole class discussions, facilitated by the lead. To aid in students' development of a more holistic model of the phenomenon, many modules incorporate outside data in addition to student-collected data. This session is also used to more formally address the CCCs, which helps tie the phenomenon to real-world applications.

2 | METHODS

2.1 | Participants and procedure

For the purposes of this study, data from 2^{nd}–5^{th} grades in the 2018–2019 school year were examined. Due to school privacy policies, the authors were not allowed access to demographic data for individual students. All nine schools were located within 15 miles of the university. The authors obtained IRB approval to collect data (IRB # 5-19-0924). 79%, 84%, 83%, and 85% of guardians consented to their student's data being used for this study in 2^{nd}, 3^{rd}, 4^{th}, and 5^{th} grades, respectively. This resulted in a sample of 173, 82, 205, and 149 students that completed the pre- and posttests in 2^{nd}, 3^{rd}, 4^{th}, and 5^{th} grades, respectively.

All assessments were administered as soon as SciTrek came to the classroom. The assessments were read aloud by the lead, and students were given approximately 3 minutes to finish. Assessments contained two questions: Things Scientists Do (TSD; "What is one/two thing(s) that scientists do, other than experiments?;" 2^{nd} and 3^{rd} grades/4^{th} and 5^{th} grades, respectively) and Attitudes ("Do you like science more, the same, or less after participating in the SciTrek program?;" posttest only). Both of these questions were chosen because of their accessibility to 2^{nd} graders, the youngest participants in our study, as well as their efficiency in delivery, as SciTrek had very limited time to assess students in the classroom. Given SciTrek's emphasis on 3D teaching, which contextualizes scientific practices in relevant NGSS-aligned phenomena, we focus on one aspect of EUS, which is an awareness of complexities within the practice-based process of scientific knowledge generation. EUS is operationally defined as a complexity score derived from coding of the students' answers to the TSD question. Students' attitudes toward science are defined in the context of our program as a potential change in students' self-reported liking of science after program participation.

Data coding for the TSD was conducted by three judges (two SciTrek staff members and one person not affiliated with SciTrek) who were trained using an older data set. Judges used a scoring rubric (created by SciTrek; Table 2) to independently code the TSD responses. Initial agreement statistics showed acceptable agreements with a single score intraclass correlation = 0.93 for all judges. The scoring allowed for a maximum of 1 point for $2^{nd}/3^{rd}$ grades (naming one thing) and a maximum of 2 points for $4^{th}/5^{th}$ grades (naming two things). Each response was scored on the most complex answer(s), using either a full point, a half-point, a quarter-point, or no point depending upon the complexity of the answer. If more than the requested number of responses were listed, the total score came from the highest scored response(s).

At the annual program completion, teachers ($N = 36$ of 41 SciTrek classroom teachers) were given a survey in which they shared their perceptions of both the program structure as well as the effects of the program on their students. They were asked about their agreement (on a 5-point Likert Scale; 1 = Strongly Disagree, to a 5 = Strongly Agree) with the following statements: 1) SciTrek improves students' understanding of scientific practices, 2) SciTrek improves students' attitudes toward science, 3) SciTrek improves

TABLE 2 Scoring rubric for Things Scientists Do (TSD)

Points awarded	Category of action	Examples
1	Epistemological processes	Make conclusions Analyze data
0.5	Engineering actions (Specific)	Creating medicine Building a bridge
0.5	Components of an experiment (specific)	Use a graduated cylinder Look at fossils
0.5	Learning processes	Study Research
0.5	Teach/collaborate	Talk about science Teach others
0.25	General subjects	Read Do chemistry
0.25	Engineering actions (general)	Build things Make stuff
0.25	Components of an experiment (general)	Lab work Take samples
0.25	Help others	Help other people Save the world
0	Daily behavior and attitudes	Work Have fun
0	Blank/irrelevant responses	Experiment Shower

Notes: Judges used the rubric to score students' responses to the open-ended question, "What is **one** thing that scientists do, other than experiments?" (2nd – 3rd grades; highest possible score 1; bold not in original) or, "What are **two** things that scientists do, other than experiments?" (4th – 5th grades; highest possible score 2; bold not in original).

students' critical thinking, 4) SciTrek integrates mathematics well, and 5) SciTrek integrates language arts well.

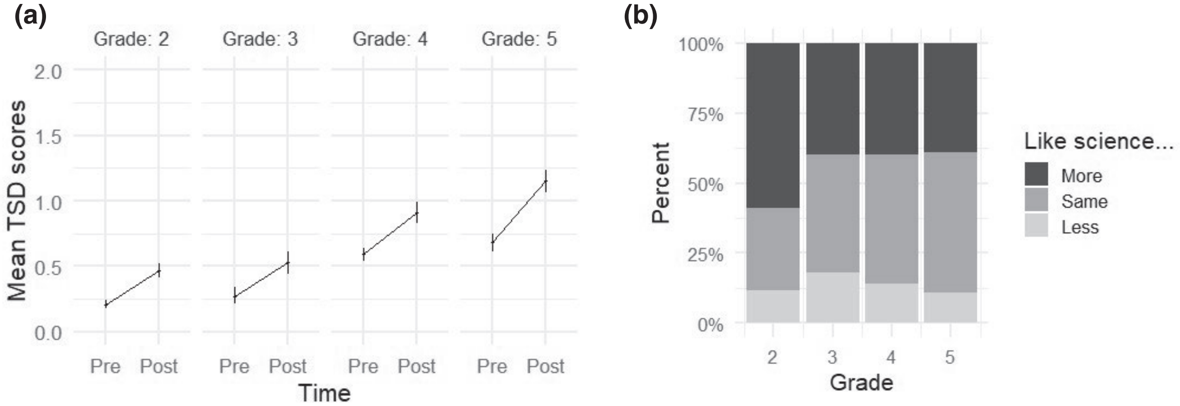
2.2 | Data analysis

All analyses were conducted using R (Version 4.0.0; R Core Team, 2020). To account for demographic differences between schools, all analyses were run as linear mixed models with random intercepts by school; no random slopes were included. Generalized linear mixed models with logistic link functions were used for analyses with binary dependent variables. Except where noted, all models were fit using the mixed() function from the *afex* package (Version 0.27–2; Singmann et al., 2020), with p-values calculated via parametric bootstrapping, and included all possible interactions of predictors. Sum-to-zero contrasts were used for all categorical predictors, allowing fixed effects to be interpreted analogously to Type III ANOVA results. Planned contrasts and post-hoc tests were conducted via the *emmeans* package (Version 1.4.6; Lenth, 2020), using the Holm method to control familywise error rates.

3 | RESULTS

3.1 | Effect of program participation on students' EUS

We conducted two linear mixed models with a 2 (time; pretest vs. posttest) × 2 (grade; 2nd vs. 3rd, or 4th vs. 5th) design predicting TSD scores (Figure 3a). Responses for 2nd and 3rd grade students were analyzed separately from those of 4th and 5th grade students due to differences in maximum possible score by grade. For 2nd and 3rd grade students, only a significant main effect of time emerged ($\chi^2 = 69.93$, $p < .001$), such that TSD scores were higher on the posttest ($M = 0.50$, 95% CI [0.45–0.55]) than on the pretest ($M = 0.24$, 95% CI [0.19–0.29]). For 4th and 5th grade students, significant main effects of time ($\chi^2 = 109.34$, $p < .001$) and grade ($\chi^2 = 6.05$, $p = .028$) emerged, qualified by a significant interaction of time and grade ($\chi^2 = 3.76$, $p = .048$). TSD scores were significantly higher on the posttest ($M_4 = 0.90$, 95% CI [0.81–1.00]; $M_5 = 1.14$, 95% CI [1.02–1.26]) than on the pretest ($M_4 = 0.58$, 95% CI [0.48–0.68]; $M_5 = 0.68$, 95%



(c)

	SciTrek Improves Students' Critical Thinking	SciTrek Integrates Math Well	SciTrek Integrates Language Arts Well	SciTrek Improves Students' Attitudes Towards Science	SciTrek Improves Students' Understanding of How Scientists Do Research
Average (SD) Rating Max: 5	4.35 (0.79)	4.24 (0.96)	3.84 (1.14)	4.19 (1.18)	4.51 (0.87)
%Teacher Ratings Somewhat/Strongly Agree	95%	78%	70%	81%	89%

FIGURE 3 Results from Things Scientists Do (TSD), attitudes, question, and post-program teacher survey. (a) Average scores of TSD (2^{nd}-3^{rd} grades have a max score of 1; 4^{th}-5^{th} grades have a max score of 2). Error bars represent bootstrapped 95% confidence intervals. (b) Percentage of students liking science more, the same, or less after participation in SciTrek by grade (Attitudes). (c) Teachers' ($N = 36$) perceptions of structure and effects of the program. Scale is based on a 5-point Likert Scale: 1 = Strongly disagree, 2 = Somewhat disagree, 3 = Neither agree nor disagree, 4 = Somewhat agree, and 5 = Strongly agree

CI [0.56–0.79]); this increase was marginally greater for 5^{th} grade than for 4^{th} grade students ($\Delta M = 0.14$, 95% CI [0.00–0.28], $p = .053$).

3.2 | Effects of program participation on students' attitudes toward science

On the posttest, 45.3% of overall students indicated that they liked science more, while 41.6% said they liked science the same and 13.1% said they liked science less. To examine how participation in the program influenced students' attitudes toward science, we conducted a generalized linear mixed model predicting the likelihood that students indicated they liked science more (versus less) by grade (2^{nd} vs. 3^{rd} vs. 4^{th} vs. 5^{th}; Figure 3b). Overall, students had a greater likelihood of indicating liking science more, compared to less (OR = 1.34, 95% CI [1.33–1.35] $p < .001$). Grade level was not observed to have a significant effect ($\chi^2 = 3.02$, $p = .450$).

3.3 | Teachers' perceptions of SciTrek structure and effects on students

On the teacher survey (post-SciTrek program), the majority of teachers rated the program very favorably (Figure 3c), with SciTrek receiving the strongest positive feedback in improving students' understanding of scientific practices, and critical thinking skills.

4 | DISCUSSION

In the current study, students generally reported liking science more after SciTrek as compared to less, which is consistent with IO and SCO program findings regarding liking of science (e.g., Clarke et al., 2019; Mooney & Laubach, 2002), highlighting the potential of SciTrek as an intervention that could impact students' attitudes toward science before the middle school years, a period that is formative for long-term retention within STEM (Krapp & Prenzel, 2011; Sadler et al., 2012). Further, this increase in students' positive attitudes toward science was also reported by teachers. Results suggest SciTrek helped students across all grade levels (2^{nd}-5^{th}) develop a deeper awareness of the complexities within the practices of scientific knowledge generation (EUS; TSD), a measure typically not assessed by IO or SCO programs. Similarly, most teachers noticed this increase in EUS, and agreed that SciTrek improved their students' critical thinking. The results from the SciTrek program are unique by recreating the benefits of traditional SCO and IO programs, in addition to incorporating several elements not found in either program type: 3D-based programming with mentors (c.f. Clarke et al., 2019), extended program support for students (multiple visits with AMKOs) and teachers (apprenticeship model of PD; c.f. Cotabish et al., 2013), and the addition of an elementary-aged appropriate EUS measure (measuring the E vs C domain; c.f. Patrick et al., 2008).

We attribute the gain in students' EUS partially to SciTrek's integrative approach, which allows students to engage in an authentic process of science alongside scientists in small groups (facilitated by AMKOs), over an extended timeframe. Further, SciTrek's focused incorporation of 3D teaching and creation of a collaborative learning environment creates a space where students take part in the process of scientific knowledge generation and are engaging in authentic scientific practices. The overall large difference in pre- to posttest scores on the TSD question supports the assertion by Furtak et al. (2012), that having students generate, develop, and justify explanations for phenomena, as they did throughout the modules, deepens their awareness of the complexities within the practices of scientific knowledge generation. At the 5^{th} grade level, a marginally greater EUS was demonstrated compared to the 4^{th} grade level. This may possibly be attributed to the process of experimental revision (5^{th} grade only), whereby students go through an iterative process of scientific knowledge generation by designing a new experiment.

A key aspect of SciTrek is the integration of grade level standards into the curriculum. The dominant standard within the module is the NGSS PE, and mathematics and language arts practices are also interwoven throughout the modules. Over two-thirds of teachers noticed this comprehensive integration and agreed that SciTrek incorporated mathematics and language arts in a way that was helpful for students.

The lack of demographic data was a limitation, and the inability to run analyses for changes in EUS and attitudes toward science among Black, Indigenous, and People of Color (BIPOC) students, students who are socioeconomically disadvantaged, and MLLs leaves important questions unanswered. BIPOC are significantly underrepresented in the science and engineering workforces (National Science Board, National Science Foundation, 2020), and assessing the impact of science education outreach programs in this regard is important. This lack of demographic data also meant that we could not use sociocultural factors to inform our data analysis. Another limitation is the use of single-measure instruments (TSD and attitudes), as these may provide a narrow view of the multifaceted nature of EUS and attitudes (Blalock et al., 2008). However, these single-measures were chosen to be easily accessible to students as young as 2^{nd} grade, and due to time constraints. In future studies, we will consider assessing the validity of our one-use measures by comparing them to validated measures of EUS and attitudes.

We would like to acknowledge that SciTrek is a resource-heavy program that might not be accessible for all classrooms. Suggestions for how to implement some of the key features of SciTrek include modifying the free source curriculum on our website to allow for: designing experiments as a class (teacher-modeled), utilizing students in upper grades to act as MKOs in lower grades, and taking part of a module (e.g., posters; as defined in Table 1) and integrating it into smaller lessons to familiarize students with that SEP before doing a complete module.

It is clear more research is needed on programs aimed at increasing science literacy in elementary students, because the majority of 5^{th} grade students in California do not achieve current science standards (California Department of Education, 2019). One way SciTrek could add to this body of knowledge is to perform an in-depth analysis of the specific SEP-like practices at each grade level to determine which aspects of the SEP students are challenged by to inform curriculum development. Other SCO and IO studies have found potential benefits in terms of gender-based effects on science identity and science interest (e.g., Farland-Smith, 2012; Patrick et al., 2008), so future work could examine whether the program has differential gender-based effects in students' interest and identities as scientists.

Overall, we have evidence that SciTrek positively affects students' EUS and attitudes toward science, a finding corroborated by their teachers. The key features of the SciTrek program can be exported to other classrooms, either as a whole or in parts, with the goal of maintaining students' interest in science to positively impact possible STEM career choices.

CONFLICT OF INTEREST

We have no known conflict of interest to disclose.

ORCID

Ryan T. Helsel https://orcid.org/0000-0001-5387-7402
Vanessa Woods https://orcid.org/0000-0002-2218-2058

REFERENCES

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. Oxford University Press.

Blalock, C. L., Lichtenstein, M. J., Owen, S., Pruski, L., Marshall, C., & Toepperwein, M. (2008). In pursuit of validity: A comprehensive review of science attitude instruments 1935–2005. *International Journal of Science Education*, *30*(7), 961–977. https://doi.org/10.1080/09500690701344578

Bybee, R. W. (2011). Scientific and engineering practices in K–12 classrooms. *Science Scope*, *35*(4), 6–11.

California Department of Education. (2019). *California assessment of student performance and Progress*. California Science Test (CAST) bit.ly/3iV8Zkz

Clarke, M. A., Sharma, N. M., & Schiller, A. M. (2019). An outreach program with hands-on, physiology-based exercises generates questions about STEM career expectations. *Advances in Physiology Education*, *43*(2), 175–179. https://doi.org/10.1152/advan.00013.2019

Cobb, P., & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist*, *31*(3-4), 175–190. https://doi.org/10.1080/00461520.1996.9653265

Cotabish, A., Dailey, D., Robinson, A., & Hughes, G. (2013). The effect of STEM intervention on elementary students' science knowledge and skills. *School Science and Mathematics*, *113*(5), 215–226. https://doi.org/10.1111/ssm.12023

Dabney, K. P., Chakraverty, D., & Tai, R. H. (2013). The association of family influence and initial interest in science. *Science Education*, *97*(3), 395–409. https://doi.org/10.1002/sce.21060

Farland-Smith, D. (2012). Development and field test of the modified draw-a-scientist test and the draw-a-scientist rubric. *School Science and Mathematics*, *112*(2), 109–116. https://doi.org/10.1111/j.1949-8594.2011.00124.x

Furtak, E. M., Seidel, T., & Iverson, H. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of Educational Research*, *82*(3), 300–329. https://doi.org/10.3102/0034654312457206

Hmelo-Silver, C. C., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and. *Educational Psychologist*, *42*(2), 99–107. https://doi.org/10.1080/00461520701263368

Krajcik, J., Codere, S., Dahsah, C., Bayer, R., & Mun, K. (2014). Planning instruction to meet the intent of the next generation science standards. *Journal of Science Teacher*, *25*(2), 157–175. https://doi.org/10.1007/s10972-014-9383-2

Krapp, A., & Prenzel, M. (2011). Research on interest in science: Theories, methods, and findings. *International Journal of Science Education*, *33*(1), 27–50. https://doi.org/10.1080/09500693.2010.518645

Lenth, R. (2020). *emmeans: Estimated marginal means, aka least-squares means*. R package version 1.4.6 http://CRAN.R-project.org/package=emmeans

Maltese, A. V., & Tai, R. H. (2009). Eyeballs in the fridge: Sources of early interest in science. *International Journal of Science Education*, *32*(5), 669–685. https://doi.org/10.1080/09500690902792385

McIntosh, K. (2003). Potato candle. *Science Scope*, *27*(1), 44–45.

Mooney, M., & Laubach, T. A. (2002). Adventure engineering: A design centered, inquiry based approach to middle grade science and mathematics education. *Journal of Engineering Education*, *91*(3), 309–318. https://doi.org/10.1002/j.2168-9830.2002.tb00708.x

Mortimer, E. F., & Scott, P. H. (2003). *Meaning making in secondary science classrooms*. Open University Press.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press. https://doi.org/10.17226/13165

National Science Foundation, National Science Board. (2020). *The State of U.S. Science and Engineering 2020*. http://ncses.nsf.gov/pubs/nsb20201/

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. The National Academies Press. https://doi.org/10.17226/18290

Patrick, H., Mantzicopoulos, P., & Samarapungavan, A. (2008). Motivation for learning science in kindergarten: Is there a gender gap and does integrated inquiry and literacy instruction make a difference. *Journal of Research in Science Teaching*, *46*(2), 166–191. https://doi.org/10.1002/tea.20276

Pedaste, M., Mäeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A. N., Kamp, E. T., Manoli, C. C., Zacharia, Z. C., & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, *14*, 47–61. https://doi.org/10.1016/j.edurev.2015.02.003

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing www.R-project.org/

Rogoff, B. (1995). Observing sociocultural activity on three plans: Participatory appropriation, guided participation and apprenticeship. In J. V. Wertsch, P. del Rio, & A. Alvarez (Eds.), *Sociocultural studies of mind* (pp. 139–164). Cambridge University Press.

Sadler, P. M., Sonnert, G., Hazari, Z., & Tai, R. (2012). Stability and volatility of STEM career interest in high school: A gender study. *Science Education*, *96*(3), 411–427. https://doi.org/10.1002/sce.21007

Singmann, H., Bolker, B., Westfall, J., Aust, F., and Ben-Shachar, M. (2020). afex: *Analysis of Factorial Experiments*. R package version 0.27-2. http://CRAN.R-project.org/package=afex

Tudge, J., & Scrimsher, S. (2003). Lev S. Vygotsky on education: A cultural-historical, interpersonal, and individual approach to development. In B. J. Zimmerman & D. H. Schunk (Eds.), *Educational psychology: A century of contributions* (pp. 207–228). Routledge.

How to cite this article: Helsel, R. T., Lambert, S., Dickerson, L., Strelich, J., Woods, V., & Feldwinn, D. (2022). Design of a phenomenon-based science outreach program and its effects on elementary students' epistemological understanding of, and attitudes toward, science. *School Science and Mathematics*, 122, 74–85. https://doi.org/10.1111/ssm.12515