Contents lists available at ScienceDirect

Journal of Applied Developmental Psychology

journal homepage: www.elsevier.com/locate/jappdp

Can service-learning boost science achievement, civic engagement, and social skills? A randomized controlled trial of Connect Science

Sara E. Rimm-Kaufman [a,*], Eileen G. Merritt [b], Candace Lapan [c], Jamie DeCoster [a], Ashley Hunt [a], Nicole Bowers [d]

[a] University of Virginia, 405 Emmet Street South, Charlottesville, VA 22903, United States of America
[b] Department of Forest Resources and Environmental Conservation, Virginia Polytechnic Institute, Cheatham Hall, 310 West Campus Drive, Blacksburg, VA 24061, United States of America
[c] Wingate University, Department of Psychology, Neu Building, 120 West Wilson Street, Wingate, NC 28174, United States of America
[d] Arizona State University, Mary Lou Fulton Teachers College, ASU West, 4701 W. Thunderbird Rd., Mail Code 3151, Glendale, AZ 85306, United States of America

ARTICLE INFO

Keywords:
Service-learning
Science education
Civic engagement
Social and emotional learning
Environmental service-learning
Energy literacy

ABSTRACT

This randomized controlled trial examined the efficacy of an elementary school service-learning program, Connect Science (CS), on classroom practices and students' science achievement, civic engagement, and social skills. Fourth grade teachers were enrolled into intervention versus control conditions resulting in 41 classrooms (20 intervention) with 868 students (423 intervention). Intervention group teachers taught CS and control group teachers taught "business-as-usual" science over a 14–22 week period. CS had a positive impact on classroom practices (Next Generation Science Standards and CS practices). Analyses showed positive impact of CS on science achievement and energy attitudes and behaviors. Use of CS practices mediated the relation between CS and energy attitudes and behaviors. Improved social skills were evident in conditions of high fidelity of implementation.

Service-learning (SL) is a form of project-based learning that meets academic learning objectives by preparing students to use their knowledge and skills to work with others to address a pressing societal problem in their community. When students use their new knowledge to work on authentic community problems, the content becomes more relevant and interesting. Such experiences promote students' engagement and learning, which in turn, helps students become better at applying their knowledge to situations beyond school (Kokotsaki et al., 2016). Current work shows promise of such whole child approaches to teaching (Celio et al., 2011; Condliffe et al., 2017). Yet, evidence-based programs that align with academic standards and support children's development across cognitive, social, and emotional domains are surprisingly rare. This scarcity is concerning given the societal challenges we face. Just acquiring knowledge in an area is not enough to motivate students to use their knowledge to work for societal change. Youth need opportunities to acquire knowledge, develop agency, and cultivate their sense of civic engagement to address important issues such as climate change now and in the future (Sanson et al., 2019).

The need for whole child approaches led us to develop and evaluate Connect Science (CS) (Harkins et al., 2019). Connect Science is a curricular approach designed to prepare teachers' to use Next Generation Science Standards (NGSS) instructional practices, teach social and collaborative skills, and facilitate a service-learning experience in their science class. The sequence of 31 lessons are typically taught over 14 weeks or more; they replace existing fourth grade science units and can be integrated with other content instruction (See https://www.connectscience.org).

CS follows eight steps described in the Appendix. In step 1, teachers create norms and teach social skills to create a sense of community and prepare for collaborative work. Then, students learn about the privileges and responsibilities that come with being an engaged citizen of their community. The class reads and discusses the book, *The Tree Lady* (Hopkins, 2013), to demonstrate how people can use their knowledge to identify needs in their community and create change. Teachers enact lessons to teach active listening, respectful communication, and respect for multiple perspectives. In step 2, teachers launch into the NGSS-aligned science lessons. Students learn about energy systems, energy production, and renewable and non-renewable resources. Lessons guide

* Corresponding author.
E-mail addresses: serk@virginia.edu (S.E. Rimm-Kaufman), egmerritt@vt.edu (E.G. Merritt), c.lapan@wingate.edu (C. Lapan), jd4nb@virginia.edu (J. DeCoster), aeh6b@virginia.edu (A. Hunt), nbowers1@asu.edu (N. Bowers).

https://doi.org/10.1016/j.appdev.2020.101236
Received 15 November 2019; Received in revised form 26 December 2020; Accepted 28 December 2020

students to discuss and debate the pros and cons of different energy sources (e.g., coal, natural gas, solar, wind, nuclear, biomass) using their newly acquired social skills. In step 3, students typically have an "ah-hah" moment when they notice the heavy reliance (~85% in the U.S.) on non-renewable resources for electricity as a problem. For instance, a student might describe the problem as, "We are using too many fossil fuels to last for our lifetime." Also, teachers assess students' science learning in this step.

Given students' new knowledge, the class works together in step 4 to propose solutions to the energy problem that they discovered in step 3. To achieve this goal, teachers use CS lessons to help students understand different ways their class can make a difference in the world, for instance, they can impact policy, educate others, or take direct action. (Common projects include energy fairs, all-school energy-use challenges, and letters to school leaders recommending a reduction in energy use.) Each class chooses how they want to take action in step 5; some classrooms chose one larger project while others chose several smaller projects, enabling youth voice in the decision-making process. Next, the class plans a project in step 6, implements the project in step 7, and assesses the impact of their project and reflects on their new knowledge and the service-learning process in step 8. Throughout the full 8-step service-learning experience, lessons focused on social and emotional skills are interspersed throughout so that students learn the skills needed to communicate effectively, work together with others on a project, reflect and make decisions, and resolve conflicts that arise. The goal of CS is to enhance students' science learning, civic engagement, and social and emotional skills.

At its core, Connect Science applies theory in developmental science to classroom instruction. CS fits with sociocultural theory that describes engaging children and youth in meaningful, authentic activities that are valued in society as essential to their development (Vygotsky, 1978). In doing so, those experiences can be transformational and lead to deep learning as well as a long-lasting understanding of how skills can be used in ways that are relevant to society (Rogoff, 1994). Science educators strive to provide elementary science instruction that is engaging, meaningful, and motivating to to students. Tethering science learning to authentic societal needs fits with educators' goals. Also, CS draws from Self-Determination Theory by creating engaging classroom experiences designed to meet students' needs for competence, autonomy and relatedness (Deci & Ryan, 1985). Effective service-learning can address these needs by providing students with opportunities to see positive consequences of their actions (supporting competence), by exercising student voice in choosing solutions (boosting autonomy), and, and by engaging in collaboration in collaboration (fostering relatedness). In turn, these psychological experiences lead students to value learning, develop confidence in their own abilities, and experience interest in learning (Deci et al., 1991). The present study is the first research on CS. We used random assignment to CS or waitlist control conditions ("business-as-usual") to examine the efficacy of CS on classroom practices and student outcomes.

What is Service-Learning?

Service-learning (SL) is "a teaching strategy wherein students learn important curricular objectives by providing service that meets authentic community needs" (Billig et al., 2005, p. 3). The term "service-learning" is often used loosely but high quality service-learning requires links to the academic curriculum, meaningful service in the community, incorporation of youth voice, reflection, meaningful community partnerships, sufficient duration and intensity and other features (National Youth Leadership Council [NYLC], 2008). Work on SL provides theoretical and, in some cases, empirical links between high quality SL and improved social skills, civic engagement, and academic learning (Billig et al., 2005; Celio et al., 2011; Center for Youth and Communities [CFYC], 2011). SL is designed to teach civic engagement, defined as the "values, beliefs, attitudes, feelings, knowledge, skills and behaviors concerned with conditions outside the immediate environment of family and friends" (Amna, 2012, p. 613). Developing civic engagement is a crucial asset for youth (Sanson et al., 2019).

Science Teaching and Learning

U.S. students lack foundational scientific knowledge and skills. In fact, fewer than 40% of fourth grade students performed at or above proficiency in science (National Assessment of Educational Progress, 2015). The U.S. is making clear efforts to improve science education, perhaps best represented by the development of the framework for K-12 science education advanced by the National Research Council. This framework led to the emergence of the Next Generation Science Standards (NGSS). The framework outlines three integrated dimensions of science learning: use of scientific and engineering practices, application of crosscutting concepts, and understanding of complex disciplinary core ideas (NRC, 2012). The standards establish new expectations for what students "should know and be able to do" as students move toward scientific literacy. By the end of twelfth grade, students should be ready to engage in public discussions on science-related issues and be critical consumers of scientific information related to their everyday lives (NRC, 2012).

Twenty states and the District of Columbia have adopted NGSS and an additional 24 states have standards based on recommendations in the framework. Given the increased prevalence of these standards, teachers need support in shifting to NGSS-aligned instruction (NGSS, 2019; Pruitt, 2014). At the broadest level, NGSS is a child-centered approach to instruction. While traditional approaches emphasize procedural skills in doing experiments, NGSS-aligned approaches focus on argumentation and consensus building after a process of investigation and analysis. While traditional instruction involves presentation of a sequence of topics in a curriculum, NGSS-aligned approaches involve students' investigation of scientific phenomena, after which students' are expected to develop coherent explanations from those investigations (Reiser, 2013). NGSS is likely to be interesting to developmental psychologists because when used well, it kindles development across cognitive, social, emotional, and even ethical domains.

Two shifts in NGSS teaching practices are most relevant for SL: eliciting and building upon students' prior knowledge, and encouraging students to explain their thinking and critique each other's ideas as they make sense of phenomena. Related to eliciting students' prior knowledge, NGSS requires teachers to reflect on real world situations that are familiar to students as a starting point for instruction. Asking students questions about their own experience with a phenomenon requires that students access information in their long term memory, which leads to deeper processing of information (National Academies of Sciences, Engineering, and Medicine [NASEM], 2018a). Gathering information on what students already know about a topic based on their experiences outside school values students' funds of knowledge, and in doing so, helps foster student engagement for students who often feel marginalized in traditional science classrooms (Calabrese Barton, Tan & Rivet, 2008). Pertaining to student critique, explanation and argument, NGSS calls for students to explain their own reasoning behind an idea and evaluate and critique the explanations of others. This focus elevates the importance of science discourse and communication (Hayes et al., 2016) and requires teachers to ask questions with more than one right answer (Colley & Windschitl, 2016).

NGSS teaching practices fit well with SL in that teachers can engage students in science learning to the point where students discover an important problem and take next steps to develop a solution. In doing so, effective teachers tap into and build upon students' prior knowledge. Students engage in critique, explanation and evaluation as they hone in on key problems and reflect on possible solutions. Pairing NGSS instruction and SL may be especially useful in fourth grade because many concepts are abstract and complex (e.g., systems thinking, energy) and SL can make these concepts meaningful and engaging. Ideally, exposure

to NGSS teaching practices in the context of SL should lead to student learning in academic, civic, and social and emotional domains.

Science Achievement

Science achievement tests provide information about student knowledge and the effectiveness of science instruction (Schneider, Krajcik, Marx, & Soloway, 2002). A review of science education programs involving 96 studies and 292 effect sizes showed an effect size of 0.41 for science interventions for grades K through 8 (J. Taylor et al., 2018). Programs vary in their efficacy; a meta-analysis based on 23 elementary school science curricula investigated 17 programs involving inquiry-oriented instruction (Slavin, Lake, Hanley, & Thurston, 2014). Seven programs with science kits (well-developed physical materials) showed an effect size of 0.02 and ten programs without kits (focused generally on processes like cooperative learning, concept development to use in science teaching) showed an effect size of 0.36 on achievement. This work provides benchmarks for expected effects.

Science knowledge provides the raw material for student problem solving in science. Further, students' prior knowledge impacts their future learning (NASEM, 2018b) and transfer of knowledge to new situations can only occur after students have acquired and consolidated content (Hattie & Donoghue, 2016). Science knowledge plays a crucial role in civic action. Without knowledge, students may choose ineffective solutions to environmental problems.

Civic Engagement

SL is designed to cultivate the civic attitudes, beliefs, knowledge and skills that develop students' ability to be change agents in society (Billig et al., 2005). In this study, we focus on four subconstructs of civic engagement that reflect affective, behavioral and cognitive states: energy attitudes, energy behaviors, civic skills, and civic efficacy. Two subconstructs (energy attitudes and behaviors) are domain-specific and focus on students' views on energy-related problems. Two subconstructs are domain-general (civic skills and civic efficacy) and pertain to students' perspective on a broad range of societal problems (Condon & Wichowsky, 2018).

As domain-specific outcomes, *energy attitudes* have an affective component and refer to sensitivity to energy issues and an awareness that saving energy is important. *Energy behaviors* are the actions and intentions related to saving energy (DeWaters & Powers, 2011). In the domain-general category, *civic skills* refers to a students' knowledge about how to create and implement a project in their community. *Civic efficacy* means that youth feel a sense of agency because they know that they can take a position on an issue that matters to them, take action, and make a difference (Serrierre, 2014).

Most SL practice and research focuses on middle and high school students (Conway et al., 2009). However, the upper elementary school years are an important time for the development of civic habits (Mitra & Serriere, 2012). Children this age become increasingly concerned about justice and fairness (Wood, 2007), setting the stage for actions that can make a difference in their community. By this point in development, children can gain knowledge in science, understand content with depth, identify a problem to address, and use knowledge to solve real-world problems (Mitra & Serriere, 2012; Rimm-Kaufman & Merritt, 2019). Engaging in collective work and exercising voice through SL can support their sense of agency, which in turn, can be a foundation for the growth of other developmental assets (Mitra & Serriere, 2012; Sherrod, 2015).

Social and Emotional Learning

Engaging in SL involves collaboration with peers as well as people in the community. Students need to be competent in their social and emotional skills to achieve these goals. Students need relationship skills involving listening, cooperation and effective management of conflict. They need social awareness involving perspective taking and empathy as they work with people whose views are different from theirs. Collaborative work can be frustrating; students need to be aware of their own thoughts and feelings and manage emotions as they work toward progress. Taking action in the community requires effective decision-making, as well (Collaborative for Academic Social and Emotional Learning, 2020).

Surprisingly few programs integrate SEL into academic instruction (Kochenderfer-Ladd & Ladd, 2016). 4Rs integrates SEL with English Language Arts (Morningside Center, 2019) and Intensified Algebra integrates SEL with math instruction (Tidd et al., 2018). This combined emphasis is particularly rare in science instruction. Yet, children's development occurs across cognitive, social, emotional and civic domains and growth in one domain impacts growth in another (Snyder & Lit, 2010). Given this reality, classroom conditions that encourage students to apply their social and emotional skills to academic work conveys certain advantages (Hunt et al., 2020). For example, teachers can leverage students' skills in meaningful discussions and collaborative work that helps consolidate learning and improves transfer of knowledge to new problems (Hattie & Donoghue, 2016).

NGSS Science and Engineering Practices (SEP) require social and emotional skills. The SEP practice, *Engaging in Argument from Evidence,* requires effective communication in that students need the ability to challenge someone about their ideas while also showing respect and perspective-taking. The SEP practice, *Planning and Carrying Out Investigations* calls for social competence including skills such as working well with peers and cooperating without prompting. Teachers who align instruction with NGSS face the challenge of preparing students with the social and emotional skills needed to attain the high quality discourse and cooperation required by the SEP (Harris & Rooks, 2010; Manz, 2015). High quality communication and social competence are essential to science-based SL.

Fidelity of Implementation

One of the major challenges in school-based research is variation in implementation. Classrooms are just one part of a complex educational system, resulting in many sources of variability in the extent to which teachers enact programs as intended. Teacher characteristics, school organizational factors, district-level influences, aspects of the program itself, and implementation support strategies (e.g., coaching) all facilitate or hinder the use of new programs (Century & Cassata, 2016; Harris et al., 2015).

In the presence of such variability, it is essential to measure fidelity of implementation (Century et al., 2010). Two themes require mention. First, research makes a strong case for measuring program-like practices in both intervention and control conditions to assess actual differences between the two groups (Abry et al., 2014; Hulleman & Cordray, 2009). Regardless of the program, many program components resemble practices used widely by typical teachers, especially given the increase in SEL in schools. For example, CS begins with establishing classroom norms. However, many teachers already begin their school year by establishing rules or norms. Teaching active listening and using sentence stems for respectful communication are key components of CS, but can show up in control classrooms, as well. Some of the CS lessons listed in the appendix resemble high quality SEL practices that teachers learn elsewhere (Denton & Kriete, 2000; Wong et al., 2009). Therefore, it is important to assess CS-like practices in intervention and control conditions. We call these "CS practices" and measure these with a custom-designed measure that can be used in both intervention and control conditions.

Second, there are a variety of possible indicators of fidelity of implementation including dosage, adherence, and program differentiation (Dane & Schneider, 1998). Program dosage gives information about the amount of program exposure but it may not be sufficient. Program adherence assesses the extent to which teachers deliver the program as designed. Not only can it be used to examine program differentiation

between groups (Century et al., 2010), it also can be more a valuable indicator of outcomes (Goncy et al., 2015). Given these issues in fidelity of implementation, we use a single measure to assess CS practices in intervention and control classrooms. We also assess dosage (percent of lessons used) in the intervention group as a secondary source of information about implementation.

Classroom Contexts and Student Individual Differences

Other teacher and student factors that may contribute to outcomes require consideration. Teachers' educational attainment and years of teaching experience may relate to outcomes (Huang & Moon, 2009). The classroom format (departmentalized versus self-contained) can affect students' learning experiences (Baroody, 2017). Student past outcomes and attributes contribute, as well. Students' initial levels of science achievement relate to later achievement (Morgan et al., 2016). Student demographics relate to academic and social inequities and need to be considered. For instance, the percentage of students living with economic disadvantages taps into experiences that may put students at a disadvantage in learning science (Morgan et al., 2016). Percent of English language learners in the classroom (NASEM, 2018a) and class size are factors that may relate to science, civic engagement and social skill outcomes. (Mihaly & McCaffrey, 2015). We include these variables as covariates.

The Present Study

This study examines the extent to which CS produced changes in classroom practices and student outcomes. Three research questions were addressed: (1) To what extent does CS impact classroom practices? (2) Do students participating in CS show improved outcomes (science achievement, civic engagement, social skills) compared to the control group? (3) To what extent does the relation between CS and student outcomes stem from changes in classroom practices (use of NGSS teaching practices, CS practices)?

Method

Participants

This study included 41 fourth grade classrooms in ethnically- and socioeconomically-diverse schools in a large public school district in South Central U.S. Fourth grade teachers were invited to participate in spring, 2017, resulting in 46 teacher participants with informed consent. Some teachers taught more than one classroom and one classroom had two teachers, thus the initial participants anticipated teaching 54 classrooms. Of the 46 participants, 22 teachers were in the intervention group and taught a total of 27 classrooms and 24 were in the waitlist control group and taught a total of 27 classrooms. Between random assignment (described below) and post-test, 8 intervention teachers left the study (which corresponded to 7 classrooms; 2 teachers who left co-taught a single classroom). Further, 6 control teachers (teaching 6 classrooms) also left. The final sample included 20 intervention classrooms (taught by 14 teachers) and 21 control classrooms (taught by 18 teachers).

The schools used grade-level teacher meetings to plan curricula. That raised concerns about intervention diffusion if teachers at the same school were assigned to different conditions. For that reason, we randomized at the school level. A methodologist who was unfamiliar with the intervention and schools randomly assigned schools to intervention versus control. Schools were stratified by quartiles using school-level fourth grade science scores from the most recent data available (2014–15) and then assigned randomly into intervention versus waitlist control groups. At baseline, *t*-tests revealed no differences between intervention and control groups on percentage of students proficient in science at the school, percentage of students at the school identified as economically disadvantaged, teacher education, teachers' years of experience, or departmental structure. However, the intervention group had a higher percentage of English Learners at the school (25%) than the control group (13%). See Table 1.

On average across groups, 23.4% of fourth graders were "proficient" in science, 49.5% of students were identified as economically disadvantaged (identified by the district as members of households receiving federal assistance [e.g., SNAP, TANF]) and 19.5% were EL. The sample was ethnically diverse with 34.4% Caucasian, 36.5% African American, 23.5% Latinx, 6% Asian, 0.6% American Indian, and 0.3% Pacific Islander.

The final sample included 32 teachers (18 intervention, 14 control) at 25 schools (12 intervention, 13 control) after 30% teacher attrition (with no differential attrition by groups). Attrition occurred if teachers transferred to teach a different grade, were no longer teaching science, or did not have enough time for CS. On average, the teachers in the final sample were 90.6% female with 9.53 (SD = 6.9) years of teaching experience. Of 32 teachers, 75% held Master's degrees, 90.6% were Caucasian and 9.4% were African American.

For science achievement and civic engagement outcomes, all students in those classrooms attending on the survey day were invited to participate (intervention $n = 423$, control $n = 445$). For social skills, six students from each classroom were selected at random resulting in 120 intervention and 124 control students. To do so, the research team acquired class rosters and selected 12 students randomly from the list using a random number generator. This list was sent to teachers who

Table 1
Comparison of Demographic Variables Between Intervention and Control Groups at Baseline.

	Mean (SD)			
Descriptive Variables	Intervention ($n = 20$)	Control ($n = 21$)	Combined ($N = 41$)	t-value ($df = 39$)
% on track for 4th grade science at the school	22.59	24.13	23.38	−0.46
	(10.51)	(11.07)	(10.69)	$p = 0.65$
% economically disadvantaged at the school	46.47	52.35	49.48	−0.97
	(16.43)	(21.85)	(19.39)	$p = 0.34$
% English language learners at the school	25.49	13.25	19.47	2.59
	(15.35)	(14.85)	(16.75)	$p = .01$
Teacher has master's degree (1 = Yes)	0.85	0.76	0.80	0.70
	(0.37)	(0.44)	(0.40)	$p = 0.49$
Teacher years of teaching experience	10.35	9.38	9.85	0.46
	(6.49)	(6.95)	(6.66)	$p = 0.65$
Departmental structure (1 = Departmentalized)	0.70	0.52	0.61	1.15
	(0.47)	(0.51)	(0.49)	$p = 0.26$

Note. Master's degree and years of teaching experience are calculated at the classroom level in this table. Some teachers teach more than one classroom. The participant section describes these variables for the final sample of teachers.

were asked to report on the first three boys and three girls on the list.

Intervention

CS was designed to introduce students to new science concepts related to energy, help students discover societal challenges related to energy use, teach students social and emotional skills needed to work with others, and support teachers to engage students in collective action to address a pressing energy-related problem in their community. CS is a professional development and curriculum that guides teachers to facilitate a SL experience with fourth graders. The instructional unit follows the eight steps in the KIDS as Planners Framework (KIDS Consortium, 2011). The SEL lessons were designed to support students' social awareness, self-awareness, self-management, relationship skills, and decision-making (CASEL, 2017). The science lessons align with NGSS disciplinary core ideas in earth and physical science (PS3 and ESS3), require students to use NGSS science and engineering practices and address crosscutting concepts (i.e., systems and system models, energy and matter [flows, cycles and conservation]).

Intervention group teachers participated in four days of professional development in Summer 2017, one day of professional development in August 2017 (after three weeks of using CS), a half-hour coaching session (with a second half-hour available upon request), on-demand support (via email) from a coach, and one 2-hour Celebration & Reflection session in November 2017. All teachers received a manual that included the 31 CS lessons, a box of science materials (e.g., batteries, bulbs, fan), and trade books to accompany lessons. Teachers implemented 30 CS lessons (each lasting between 30 and 50 min) distributed over a 14–22 week period starting in August 2017. CS lessons replaced three weeks of science lessons and teachers integrated remaining lessons into other subjects, resulting in a total dosage of 900 to 1500 min.

Control group teachers used local curricula to meet the same science standards. The existing science curricular unit focused instruction on analyzing how different natural resources are used to solve human problems or improve the quality of life. The activities focused on reading informational texts about natural resources and their impact on the environment, identifying and explaining properties of earth materials, exploring how power is created in their local community, and discussing how that production impacts the environment. One activity was to discuss and write an opinion piece about resource use in the community. To achieve the curricular goals, control group teachers used materials from two kits, one with materials pertaining to electric circuits and the other with materials on natural resources. Control group teachers received CS training and materials in summer, 2018, after data collection was complete.

Procedures

Data were gathered from three sources: (a) school-level district data, (b) teacher-report surveys, and (c) student-report surveys. Prior to randomization, school-level district data were collected on: 1) %-age of students at each school on track for 4th grade science, 2) %-age economically disadvantaged, and 3) %-age English learners.

Window 1 occurred within 2 weeks of the summer professional development training to gather teacher baseline data (e.g., teacher education, years of experience, departmental structure). Subsequent windows corresponded to specific benchmarks in teachers' science sequence to account for the fact that teachers varied in their pace through CS and the science kits. To ensure precise timing of data collection, the project manager communicated individually with each teacher to benchmark his/her progress through the science sequence. At certain benchmarks, the project manager distributed surveys and noted how many weeks had elapsed from the start of the science unit to define data collection windows.

Following this procedure, window 2 occurred at the end of CS Step 2 for CS teachers, after the class had completed all or most of the science-focused CS lessons but before the class launched into the SL projects. See Appendix. Window 2 occurred at a comparable milestone in the science curriculum for control teachers, after most of the science lessons from the energy kit were complete. Thus, window 2 occurred between 10 and 13 weeks after each teacher's start of their science unit and depended on each teacher's pace through the units. In window 2, teachers reported on NGSS instructional practices used since they began their units.

Window 3 occurred at the completion of the SL projects in intervention groups and upon the completion of the complete energy unit in control groups. Again, the project manager followed the same procedures as described above to link the timing of data collection to the teachers' completion of the instructional unit. At the end of the units, the project manager sent paper surveys to measure students' science achievement and civic engagement and a cover sheet with questions for teachers to complete about the classroom demographics (e.g., number of EL students, class size) to be used as potential covariates. The project manager also distributed electronic surveys to teachers for teachers to report on their use of CS practices and student social skills (for six randomly selected students from each class). Thus, benchmarked to the start of the CS and science units, window 3 student data were gathered between 12 and 18 weeks after the unit start and window 3 teacher data between 14 and 22 weeks after the unit start.

Measures

District data

District data was conveyed from the Office for Research and Evaluation to the research team electronically. Test score data was based on the state standardized test scores in science and indicated the percent of fourth grade students on track in science in 2014–15 (i.e., the latest data available). School-level demographics were ascertained from district data resulting in values indicating percentage of students identified as economically disadvantaged, percentage of English learners (EL), and student ethnicity at each school.

Teacher information surveys

Teachers reported on gender, ethnicity, education, years of teaching experience, and whether they taught departmental vs. self-contained classrooms in window 1. Teachers provided classrooms information (e.g., percent of EL, class size) in window 3 at the time of the student-report data collection to give us more analytical precision, especially given the baseline differences in EL at the school level. *T*-tests on these variables confirmed that there were more EL students in the intervention than control group, $t(39) = 2.15$, $p = .04$, but showed no difference for class size, $t(39) = -0.33$, $p = .97$.

Classroom practices

In window 2, teachers reported on *NGSS instructional practices*: (a) *eliciting and building upon students' prior knowledge* (3 items), and (b) *facilitating student critique, explanation and argument* (5 items; Hayes et al., 2016). Questions about eliciting students' prior knowledge included, "How often do you apply science concepts to explain natural events or real-world situations?" Questions about facilitating student critique, explanation and argument included, "How often do your students supply evidence to support a claim or explanation?" Teachers responded on a 1 to 5 scale (1 = never to 5 = daily or almost daily). (Alpha = 0.82 and 0.87, respectively.)

In window 3, teachers reported on *CS practices* (or CS-like practices in control schools). Teachers in intervention and control conditions reported on the practices they used since the start of the unit using a 12-item measure adapted from the SL Classroom Activity Report (KIDS Consortium & CFYC, 2008) and piloted prior to use. The measured was designed to assess program adherence, one aspect of fidelity of

implementation. It was conducted in both intervention and control classrooms in case teachers in the control group were using practices that resembled CS. Example practices included: "students practiced strategies for active listening," and "students implemented a service-learning project." Response options ranged from 1 = not at all true to 4 = very true (alpha = 0.92).

To validate the CS practices measure, intervention dosage was measured by asking teachers in the intervention group to complete lesson logs via a semi-monthly electronic survey for the full duration of CS. The survey asked teachers if and how they used the lesson (e.g., How closely did you follow the lesson versus adapt it?). Response options were reduced to create a dichotomous variable and a mean across the 30 items was computed (mean = 0.91; range = 0.70–1.0) for each teacher to create a variable indicating the percent of lessons used. On average, intervention teachers used 91% of the lessons, indicating high usability. The percent of lessons used correlated with the CS practices measure (r = 0.53, $p < .02$), despite the limited range in the CS practices measure (range = 2.58 to 4 on a 1 to 4 scale) and differences in the meaning of dosage and adherence.

Science achievement

A 13-item student survey measured students' energy knowledge as an indicator of *science achievement.* The paper survey included multiple choice questions about electricity (2 items), energy forms and systems (4 items), and energy sources (7 items) (alpha = 0.68). Despite thorough review, we were unable to find an openly available measure of student achievement related to energy concepts for this grade level. Two science education experts (one faculty and one doctoral student with cumulative experience of 32 years of science teaching) developed the measure. The team reviewed existing state, national and international assessments and compiled existing open-source multiple-choice questions related to the constructs. Next, the team created a survey using some existing, modified, and new questions, and asked six experts to evaluate test items for clarity and alignment with district fourth grade learning goals and CS content. The team conducted a think-aloud with three upper elementary school students; students took the assessment and provided comments about item clarity and difficulty. The team selected a subset of items based on three criteria: (a) clarity (to experts and students), (b) close alignment with targeted fourth grade science standards, and (c) content overlap with district curriculum and CS.

Civic engagement

Students reported on civic engagement as *energy attitudes and behaviors* (domain-specific) and *civic skills and efficacy* (domain-general). Students reported on *energy attitudes and behaviors* using an 8-item scale. No comparable scales existed for upper elementary students, thus a measure designed for eighth grade students was adapted by science education experts (Bodzin et al., 2013) and piloted with 144 fourth graders prior to this study. Example items include: "I would do more to save energy if I knew how" (energy attitude) with response options of 1 = strongly disagree to 5 = strongly agree and "When I leave a room, I turn off the lights" (energy behavior) with response items of 1 = never to 5 = always. The alpha value was 0.65; possibly lower than typical standards because of use of student-report data.

Students reported on their *civic skills* using a three-item measure asking questions such as "How well can you identify needs or problems that are important in your school or community" which they rated on a four-point scale (1 = not at all to 4 = very well) (CYFC, 2011). To measure *civic efficacy*, students completed a four-item measure with questions such as, "I can make a difference in my community" and "Students my age can do things to make the world better" which they rated from 1 = strongly disagree to 4 = strongly agree (Caswell et al., 2011). The two scales were highly collinear; scale validation analyses showed better model fit by treating them as a single scale, (alpha = 0.67). These measures were piloted before use, as well.

Confirmatory factor analyses (CFA) were conducted to examine whether to treat civic engagement as a unitary construct or two separate subconstructs. The CFA results were comparable for a one- and two-factor solution but the two-factor approach had stronger conceptual fit. The two variables, energy attitudes and behaviors and civic skills and

Table 2
Descriptive Statistics by Condition for Classroom Practices and Student Outcomes.

	Mean (SD)			*t*-test	Effect size *d*
Classroom Practice Outcomes	Intervention (n = 20)	Control (n = 21)	Combined (N = 41)		
Eliciting & building on prior knowledge	4.23 (0.45)	3.89 (0.61)	4.06 (0.56)	$t(39) = 2.06$	0.62
Student critique, explanation & argument	4.20 (0.44)	3.69 (0.75)	3.94 (0.66)	$t(39) = 2.67^{*}$	0.81
Connect Science practices	3.50 (0.43)	2.55 (0.66)	3.02 (0.73)	$t(39) = 5.45^{**}$	1.66
Science & Civic Engagement Outcomes	Intervention (n = 423)	Control (n = 445)	Combined (N = 868)		
Science achievement	0.79 (0.18)	0.72 (0.18)	0.75 (0.18)	$t(712) = 5.27^{**}$	0.39
Energy attitudes and behaviors	3.97 (0.55)	3.82 (0.62)	3.88 (0.59)	$t(682) = 3.26^{**}$	0.26
Civic efficacy and skills	3.32 (0.30)	3.24 (0.50)	3.27 (0.46)	$t(687) = 2.37^{*}$	0.19
Social Skill Outcomes	Intervention (n = 120)	Control (n = 124)	Combined (N = 244)		
Communication	3.61 (0.43)	3.55 (0.58)	3.58 (0.51)	$t(242) = 0.37$	0.12
Social competence	3.27 (0.72)	3.28 (0.72)	3.28 (0.72)	$t(242) = -0.15$	−0.01

Note. NGSS practices (eliciting and building on prior knowledge and student critique, explanation and argument) were gathered in Window 2. All other data in this table were gathered in Window 3. Positive t values and effect sizes indicate higher values in the intervention than control group.

efficacy, were correlated 0.63, suggesting considerable overlap but still distinctions worth making given the exploratory nature of the study.

Social skills

Teachers reported on social skills for six students. The measure included seven items on *communication skills* using the Social Skills Improvement System (Gresham & Elliott, 2008). Statements included, "takes turns in conversations"; teachers rated each item on a scale from 1 = never to 4 = very often (alpha = 0.88). Teachers reported on *social competence* using a six-item measure developed by Child Trends (2014). Statements included, "worked well with peers" and values ranged from 1 = none of the time to 4 = all of the time, (alpha = 0.97). CFA analyses were conducted to consider whether to treat social skills as one or two factors; the two-factor solution was a better fit than the one-factor solution and also matched the initial measures and hypothesis.

Results

Descriptive statistics

Table 2 reports means and standard deviations for classroom practices and student outcomes aggregated to the classroom level. These *t*-test analyses show higher values on classroom practice outcomes, science achievement and civic engagement outcomes in the intervention than control condition but no statistically significant difference in social skill outcomes between intervention and control conditions. These analyses are preliminary given that they do not account for covariates or adjust for the nested structure of the data.

Bivariate relations among outcomes

Table 3 shows the correlations for all variables. The three measures of classroom practices are all moderately correlated (ranging from r = 0.35 to 0.67). Within the student outcomes, energy attitudes and behaviors were correlated significantly with civic efficacy and skills (r = 0.63), and communication was correlated to social competence (r = 0.75). Both assignment to intervention condition and use of CS practices correlated significantly with three of the five student outcomes (science achievement, energy attitudes and behaviors, and civic efficacy and skills), but not social skills.

Missing data analysis

To prepare for the predictive models, we analyzed missing data at the classroom and student level. At the classroom level, all classrooms had complete teacher-report data. We had complete data at 38 of 41 classrooms (92.7%). Science achievement data was missing from one classroom (2.4%) and energy attitudes and behaviors and civic efficacy and skills data were missing from two classrooms (4.9%). Analyses did not reveal any indication that the data were not missing at random.

At the student level, two-thirds of the variables and 70% of the students had no missing data. The student-report measures were missing approximately 20% of their data (17.7% for science achievement, 20.6% for efficacy, 21.2% for energy attitudes and behaviors). Among these three variables with substantial missing data, missingness was somewhat higher in the treatment condition (mean of 24%) than in the control condition (mean of 17%). There was little missing data on social skills (< 1%). Analyses suggested that data were missing at random. There was no reason to believe that the reason for missing data/attrition related to the intervention. The study has low attrition and low expected bias according to the What Works Clearinghouse (WWC) attrition standards (WWC, 2015).

Table 3
Correlations among study variables.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Condition (1 = CS intervention)	-															
Classroom Practices																
2. Building on Prior Knowledge	0.35**	-														
3. Student Critique/Explanation	0.43**	0.66**	-													
4. CS Practices	0.67**	0.46*	0.40**	-												
Student Outcomes																
5. Science Achievement	0.19**	0.12*	0.11**	0.25**	-											
6. Energy Attitudes & Behaviors	0.12**	0.01	0.03	0.14**	0.08*	-										
7. Civic Efficacy & Skills	0.09*	0.00	0.06	0.11**	0.10*	0.63**	-									
8. Communication Skills	0.06	0.02	−0.03	0.02	0.10	0.19**	0.11	-								
9. Social Competence	−0.01	0.06	0.01	−0.01	0.14	0.26**	0.21**	0.75**	-							
Covariates																
10. Teacher Education Attainment	0.09**	−0.26**	−0.17**	−0.12**	0.02	0.07	0.08*	−0.06	−0.04	-						
11. Years of Teaching Experience	0.10**	0.05	0.02	0.12**	0.21**	0.08*	0.06	0.06	0.11	−0.10**	-					
12. Departmental Organization	0.18**	−0.29**	−0.19**	0.01	0.02	0.01	0.01	−0.11	−0.21**	0.51**	−0.12**	-				
13. Student Science Achievement	−0.08*	0.08*	0.08*	0.10**	0.23**	−0.02	−0.01	0.19**	0.10	−0.19**	0.11**	−0.03	-			
14. % Economic Disadvantage	−0.16**	0.02	−0.22**	−0.25**	−0.32**	−0.02	−0.05	−0.21**	−0.13*	0.09**	−0.14**	−0.08*	−0.79**	-		
15. % English Language Learners	0.26**	−0.20**	−0.22**	0.03	−0.09*	0.11**	0.07	−0.05	−0.04	0.23**	−0.03	0.33**	−0.47**	0.32**	-	
16. Class Size	0.03	−0.26**	−0.19**	0.06	0.15**	0.10*	0.07	−0.07	−0.03	−0.07*	0.05	0.12**	0.16**	−0.17**	0.06	-

Note. For Departmentalized Organization, 1 = departmentalized. * $p < .05$. ** $p < .01$.

Predictive modeling strategy

All predictive models were estimated using Mplus version 8. Coefficients were obtained using full information maximum likelihood (FIML) estimation. FIML accounts for missing data by using all available data for each case to estimate parameters and adjust for potential bias in the estimates resulting from missing data. This approach has been identified as one of the optimal ways to handle missing data (Peugh & Enders, 2004).

The analytic models make assumptions about the residuals: normality, independence, and equality of variance for different levels of the predictors. We tested the validity of these assumptions through residuals plots, histograms, and scatterplots. Although the residuals for communication and social competence showed evidence of skewness, none of the residual distributions were sufficiently non-normal to warrant concern about the assumption of normality. Plots of the residuals against model predictors did not reveal violations of the assumption of equal variances. Nor did these plots show evidence of non-independence caused by nonlinear relations between the predictors and outcomes. No outliers were apparent.

Although the randomization was conducted at the school level because of potential contamination, there were only seven cases where there was more than one teacher per school. Analyses were conducted at the classroom level and accounted for students clustered in classrooms. This approach maximized power and fit with early-stage research on a new program.

Research question 1

The first question examined the impact of CS on classroom practices (NGSS instructional practices and CS practices indicating use of CS-like practices/fidelity of implementation). Separate regression analyses were conducted for each of the three dependent variables: (a) eliciting and building prior knowledge, (b) student critique, explanation and argument, and (c) CS practices. Independent variables include condition (intervention = 1) and covariates (i.e., teacher education, years of experience, and departmental structure; percentage of students on track for fourth grade science, economically disadvantaged at the school; percent EL in the classroom and class size). Data were analyzed at the classroom level.

Results indicate that CS had significant positive effects on all three measures of classroom practices, as shown in Table 4. The magnitude of the effects based on the standardized betas were 0.55 and 0.49 for eliciting prior knowledge and student critique, explanation and argument, respectively, and 0.69 for CS practices.

Research question 2

The second question examined main effects of CS by examining whether CS students showed improved outcomes compared to the control group. Five regression analyses were conducted. Each outcome was examined in a separate model that controlled for covariates.

Intraclass correlations (ICCs) were examined to determine the need to account for the nested structure of the data. The ICC values showed more variability at the student than classroom level. ICCs were small for energy attitudes and behaviors and civic skills and efficacy (0.06, 0.05, respectively) but larger for science achievement (0.22), communication (0.26) and social competence (0.21). Given the nested structure of the data, all analyses with student outcomes accounted for clustering (students within classrooms). We used a sandwich estimator to adjust the standard errors of the coefficients to account for clustering at the classroom level using the "Type = Complex" command in MPlus (Muthén & Muthén, 1998–2010). Using sandwich estimators has advantages over traditional multi-level models when the number of clusters is relatively small (i.e., 41 classrooms) and there are small cluster sizes ($n = 6$ for social skills outcomes) (McNeish et al., 2017).

Findings from the regression analyses showed that CS had statistically significant positive effects on science achievement and energy attitudes and behaviors. Based on the standardized betas, we can estimate the size of the effect to be 0.32 and 0.31 standard deviations, respectively. See Table 5 for results. Although simple *t*-tests (Table 2) showed higher civic efficacy and skills in intervention than control condition, the more conservative and precise analysis that adjusts for clustering and includes covariates reveals no statistically significant main effects for civic efficacy and skills. Findings showed no impact of CS on social skills.

Table 4
Effects of Intervention on Classroom Practices.

	Eliciting & Building on Prior Knowledge			Student Critique, Explanation & Argument			Connect Science Practices		
	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*
Intervention Effect									
Connect Science	**0.55**	**0.12**	**<0.001**	**0.49**	**0.12**	**<0.001**	**0.69**	**0.09**	**<0.001**
Covariates									
Teacher has masters	−0.17	0.14	0.22	−0.11	0.15	0.46	−0.12	0.13	0.34
Years of teaching Experience	−0.02	0.10	0.83	−0.07	0.12	0.57	0.01	0.10	0.93
Department Structure	−0.16	0.18	0.35	−0.15	0.18	0.39	−0.04	0.15	0.78
% on track for 4th grade science	0.35	0.19	0.07	−0.20	0.26	0.46	0.03	0.17	0.87
% economically disadvantaged	**0.40**	**0.18**	**0.03**	−0.25	0.25	0.30	−0.09	0.16	0.59
%EL at classroom level	−0.19	0.16	0.24	−0.27	0.15	0.07	−0.06	0.13	0.66
Class Size	**−0.23**	**0.12**	**0.03**	−0.18	0.10	0.08	0.04	0.12	0.75
Overall R^2		0.51			0.40			0.42	

Note. For departmental organization, 1 = departmentalized. For Connect Science, 1 = intervention condition. Statistically significant findings are in bold.

Table 5
Effects of Intervention on Student Outcomes.

	Science Achievement			Energy Attitudes & Behaviors			Civic Efficacy & Skills			Communication			Social Competence		
	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*
Intervention Effect															
Connect Science	**0.32**	**(0.10)**	**<0.01**	**0.31**	**(0.12)**	**0.01**	0.21	(0.12)	0.07	0.06	(0.18)	0.73	−0.06	(0.18)	0.76
Covariates															
Teacher has masters	0.07	(0.12)	0.58	0.24	(0.13)	0.07	**0.26**	**(0.11)**	**0.02**	0.03	(0.12)	0.81	0.18	(0.11)	0.10
Teaching experience	0.30	(0.15)	0.06	0.11	(0.12)	0.39	0.07	(0.11)	0.50	0.03	(0.15)	0.83	0.13	(0.12)	0.27
Department Structure	−0.07	(0.15)	0.66	**−0.40**	**(0.12)**	**<0.01**	**−0.32**	**(0.15)**	**0.04**	−0.22	(0.19)	. 25	**−0.49**	**(0.17)**	**<0.01**
% on track for 4th grade science	−0.07	(0.20)	0.73	0.30	(0.20)	0.12	0.09	(0.25)	0.70	0.13	(0.17)	0.47	−0.06	(0.15)	0.70
Percent econ. Disadvantaged	**−0.56**	**(0.18)**	**<0.01**	0.01	(0.17)	0.98	−0.21	(0.22)	0.36	−0.31	(0.20)	0.12	**−0.34**	**(0.17)**	**0.04**
Percent EL	0.01	(0.12)	0.99	**0.57**	**(0.13)**	**<0.001**	0.46	(0.17)	**<0.01**	0.07	(0.19)	0.70	0.10	(0.18)	0.58
Class Size	0.08	(0.13)	0.54	−0.08	(0.15)	0.60	−0.14	(0.14)	0.31	−0.19	(0.12)	0.11	−0.04	(0.11)	0.73
Overall R^2		0.57			0.52			0.48			0.20			0.25	

Note. Department structure is dummy coded where 1 = departmentalized. Connect Science is dummy coded where 1 = intervention condition. Statistically significant findings are in bold.

Table 6
Connect Science Practices as a Mediator of the Relation between Connect Science and Outcomes.

	Science Achievement			Energy Attitudes & Behaviors		
Model Fit Indices						
χ^2 / *df*	3.13 / 7			3.13 / 7		
RMSEA [90% C.I.]	0.00 [0.00, 0.10]			0.00 [0.00, 0.10]		
CFI	1.00			1.00		
TLI	1.25			1.25		
SRMR	0.07			0.08		
R^2	0.59			0.56		
Direct Effects	*B* (SE)	*β*	*p*	*B* (SE)	*β*	*p*
Condition ➔ CS Practices	**0.95 (0.20)**	**0.67**	**<0.001**	**0.95 (0.20)**	**0.67**	**<0.001**
CS Practices ➔ Outcome	0.03 (0.02)	0.23	0.08	**0.09 (0.04)**	**0.26**	**0.02**
Condition ➔ Outcome	0.03 (0.02)	0.16	0.11	0.07 (0.07)	0.13	0.65
Indirect Effect	*B* (SE)	*β*	*p*	*B* (SE)	*β*	*p*
Condition ➔ CS Practices ➔ Outcome	0.03	0.15	0.11	**0.09 (0.04)**	**0.17**	**0.02**

Note. Covariates were included in each mediation model, but are excluded from this table for ease of interpretation. The significant paths are in bold.

Research question 3

The third question examined the extent to which the relation between CS and student outcomes stem from changes in classroom practices. In cases where a main effect of CS on student outcomes was established, mediation analyses were conducted to examine the extent to which classroom practices explained the relation between CS and student outcomes. Six mediation analyses were conducted. We tested three classroom practice variables as potential mediators: (1) eliciting and building on prior knowledge, (2) student critique, explanation and argument, and (3) CS practices. Each mediator was tested for two outcomes: (1) science achievement, and (2) energy attitudes and behaviors. We estimated the indirect effects of the intervention on each student outcome through each classroom practice using normal-theory tests.

Separate models were used to examine each indirect effect and all models included the covariates in Table 3. Sandwich estimators were used to adjust the standard errors. Results revealed that CS practices significantly mediated the effects of the intervention on energy attitudes and behaviors. In the model testing the indirect effect, the direct effect of treatment condition on energy attitudes and behavior was no longer significant indicating that CS practices fully mediated the relation between CS and energy attitudes and behaviors, as shown in Table 6. None of the remaining models showed mediation.

Moderation analyses

Since no main effects were evident for communication and social competence, we explored possible moderation. CS practices was treated as a moderator of the relation between CS and communication and social competence to examine if CS related to social skill outcomes only in high fidelity conditions. We hypothesized that the contribution of the intervention would be stronger if teachers engaged in more CS practices.

The results indicate that the effects of the intervention on communication are significantly moderated by CS practices (B = 0.42 [SE = 0.22], $\beta = 0.38$, $p < .05$), while the effects of the intervention on social competence moderated by CS practices were marginally significant (B = 0.53 [SE = 0.32], $\beta = 0.24$, $p = .07$). See Fig. 1. CS relates to higher social skills when classrooms have strong implementation, but lower social skills when classrooms have weak implementation. The reverse is true in control classrooms.

We gathered data on dosage by asking intervention group teachers to log their CS lessons as a way to validate the measure of CS practices. We furthered the exploration of fidelity of implementation in the intervention group by testing the main effects of dosage on the five student outcomes while controlling for covariates. Results showed positive

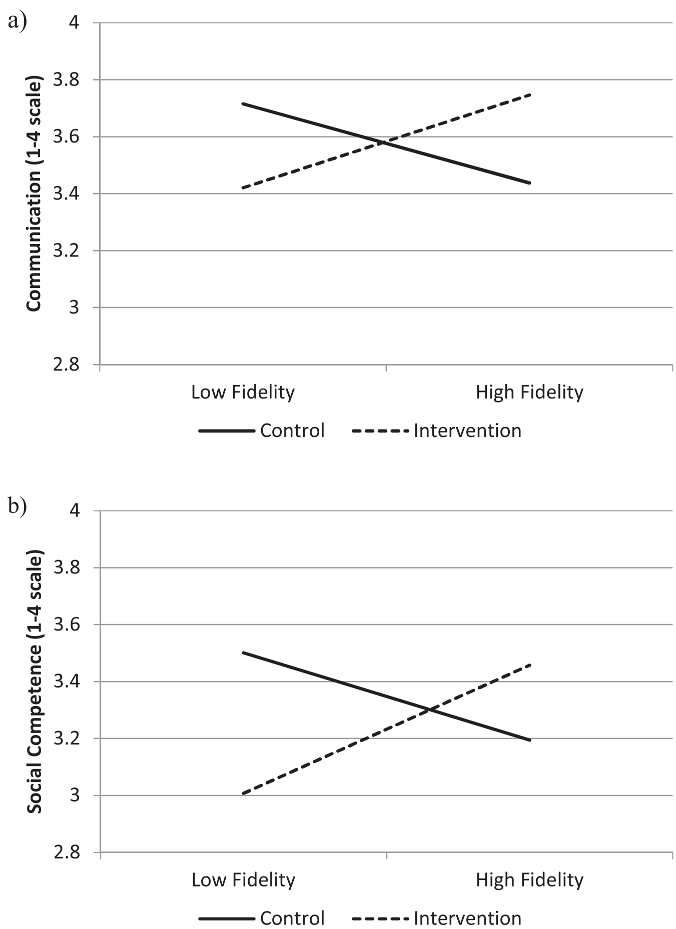

Fig. 1. Relation between Connect Science and social skill outcomes in conditions of high and low use of Connect Science practices (fidelity of implementation). Fig. 1a shows results for communication and Fig. 1b shows findings for social competence.

associations for two of the five outcomes. Teachers who implemented a higher percent of lessons had students with higher energy attitudes and behaviors ($B = 0.01$ [SE = 0.00], $\beta = 0.30$, $p < .05$) and higher social competence ($B = 0.02$ [SE = 0.01], $\beta = 0.64$, $p < .01$), but no association for the other outcomes.

Intervention description

A description of an actual fourth grade classroom brings to life the CS steps and instructional experiences listed in the Appendix. In step 1 and 2, the class began CS with a set of lessons that provided them with background knowledge in important science and SEL skills. At the end of step 3, the class chose a problem to work on; in this case, the class identified wasting electricity as a problem in their community. Next, in step 4, students began to learn civic skills as they considered solutions to their chosen problem. Fig. 2a shows a chart describing three approaches to solving problems: taking direct action, educating others or working to influence policy. Students generated ideas and the class categorized those ideas by each approach. In step 5, they used a decision-making matrix to choose a project that would have a positive impact on the problem, be feasible, and also felt important to the students in the class. In step 6, each student summarized their plan on a project description worksheet, which articulated the problem and solution in students' own words (see Fig. 2b).

Step 7 involved implementing the plan. This class decided to educate other students and their families by hosting an energy festival as their solution to the energy problem they discovered. ("Energy festival" was the fourth graders' language for an energy fair.) Students worked in small groups on booths for the festival. According to the teacher, "students created a booth to teach why we need to conserve, a demonstration both for how to conserve, a station to make magnets/stickers to promote conserving electricity, a song about conservation that they taught the student visitors, a fishing for facts game, and a bowling for information game. Adults were able to complete a home energy audit that we found on our local power company's website. After completing the audit participants registered to receive an energy saving kit. Students gave a survey to visitors at the end of the energy fair to measure what people learned as an indicator of impact. In response to "How will we know if we achieved it?" the student responded, "We will know by taking a survey on how it changed."

In Step 8, the class engaged in reflection and evaluated the impact of the energy festival on their chosen energy problem. One student said, "when you're in a small group and something went wrong all you have to do is some respectful communication." Another student reflected on their learning about electricity as well as SEL. "My family saves electricity now and so do I… When I did the energy project that probably was the most time I have interacted with peers. I got along very well. I planned it out and learned how to plan a big project." The example here shows one of many paths that a class took in their design and enactment of SL. SL allows each student and class to have some autonomy in taking action on problems that matter to them and using their own solutions ideas.

Discussion

Four main findings emerged. First, CS produced sizeable impacts on fourth grade teachers' classroom practices. Teachers trained in CS reported more NGSS-aligned classroom practices and higher levels of CS practices than control group teachers. Second, students participating in CS showed higher science achievement and boosts in domain-specific (students' attitudes and behaviors) but not domain-general civic engagement (civic skills and efficacy) compared to the "business-as-usual" condition. Third, actual use of CS practices (i.e., use of sentence stems in classroom discussions, investigation of a problem in their community, implementation a SL project) emerged as an explanatory mechanism of the relation between CS and domain-specific civic engagement (energy attitudes and behavior). No other mediation effects were statistically significant. Fourth, contrary to expectations, there were no main effects of CS on social skills. However, analyses revealed the importance of fidelity of implementation; CS related to higher communication and social competence in conditions of high implementation.

Recent trends suggest the need for integrative science curriculum that support whole child development. We see some convergence among science educators, school leaders, and teachers signaling the importance of science that takes a project-based approach. For example, science teachers are prioritizing curricula that support teachers' use of NGSS practices, emphasizes relevance to real world issues, and prioritizes authentic audiences for student work (Lieberman & Seydel, 2019; Vieira & Tenrerio-Vieira, 2016). There has been increased interest in SEL instruction; almost three-quarters of elementary principals place high priority on social and emotional based on a national survey of more than 3500 principals (Hamilton et al., 2019). The most pressing reason to shift toward integrative science curriculum comes from the needs of children and youth themselves – climate change is producing anxiety among children and programs that help students take action in their community hold promise for mitigating stress (Sanson et al., 2019). Taken together, current trends point to a reshaping of science education toward integrated approaches that situate the child as a developing learner who actively makes sense of science (Linn et al., 2016). However, just developing integrative approaches is insufficient. The field also needs rigorous research examining the efficacy of such approaches used in real-world conditions (Furco & Root, 2010). The present study

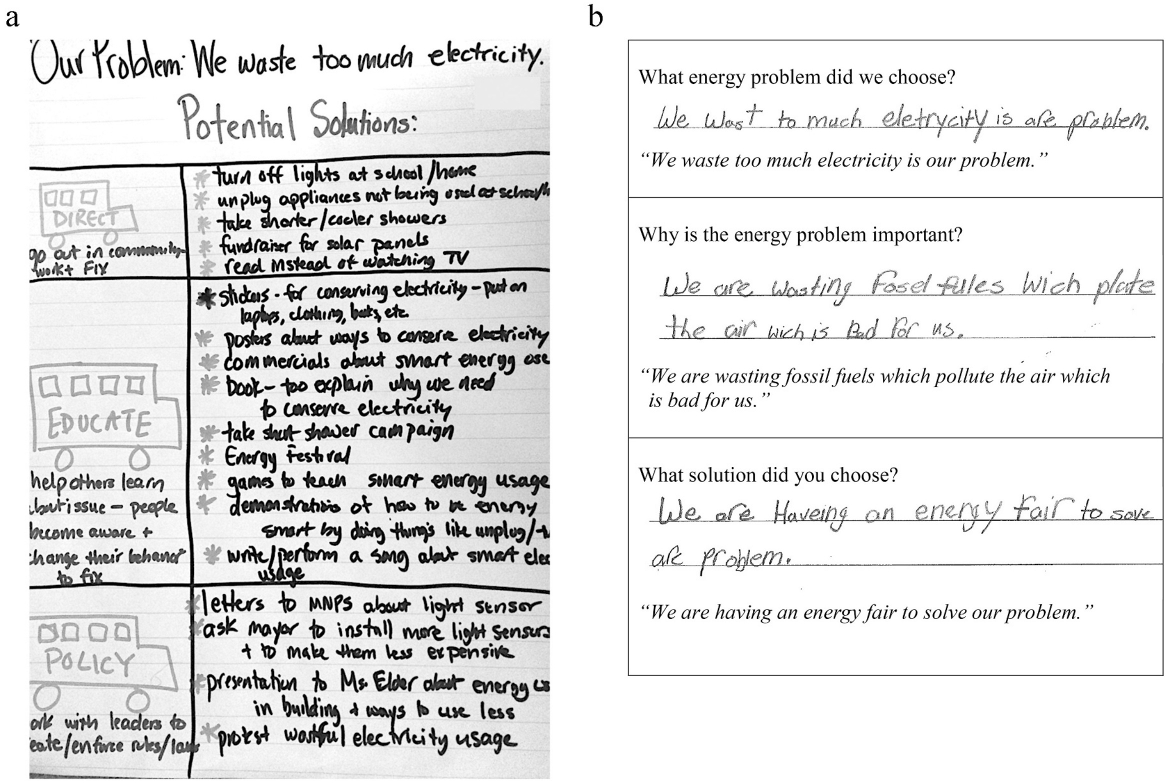

Fig. 2. 2a) Students learned three types of solutions to problems. Engaged citizens can take direct action, educate others, or influence policy. First, the teacher defined these different approaches. Then, the teacher led a brainstorming session to generate solutions that fit into these three categories. 2b) Each student wrote a project description to deepen their understanding and give the teacher feedback on how well students understood the purpose of the project.

matches this critical need.

Impact on classroom practices

Results showed a sizeable impact of CS professional development and materials on NGSS instructional practices including elicitation of students' prior knowledge and student critique, explanation and argument. Ample research in education points to the importance of accessing students' prior knowledge because it reduces attentional demands during instruction and aids teachers in their efforts to identify representations of ideas that can interfere with future learning (NASEM, 2018). The findings match hypotheses and fit with the design of CS. In relation to prior knowledge, every lesson begins with instructional strategies to link to students' prior understanding. For example, in one lesson, teachers ask students about family members who work at power plants or in the energy industry as a way of making very abstract content more relevant. Understanding scientific phenomena from the students' perspective is emphasized in CS professional development. For example, the professional development offers teachers hands-on experiences with energy-related content so that teachers can experience lessons as learners themselves. This experience, in turn, is intended to help teachers ask good questions of their students.

The curricular materials incorporate tasks that engender critical thinking, offer evidence enabling students to devise claims, give explanations, and argue for certain positions. One exercise involves a chart with pros and cons for eight energy sources. Students need to consider the advantages and disadvantages of each energy source and sort flashcards in order from best to worst sources for the future. Students work in pairs and teachers encourage students to back their claims with evidence. NGSS-aligned tasks serve as a foundation for science teachers' to respond well to students emerging understanding of concepts, facilitation of effective student discourse, and allowance of student agency (Hunt et al., 2020).

The CS examples are consistent with the ways that NGSS science and engineering practices shift away from teacher- toward student-centered instruction. Teachers need scientific knowledge and confidence in that knowledge to ask questions that tap into students' prior understanding of concepts. Teachers need guided practice supporting students' construction of explanations and students' use of claims in arguments given that teachers often struggle with this skill (Davis et al., 2017). Educative curriculum materials build teacher knowledge and pedagogical skills that are critical for teachers to engage in student-centered instruction (Davis et al., 2017; Roseman et al., 2017).

Findings demonstrate a sizeable impact of CS on use of Connect Science practices. Knowing that it can be difficult to change teacher behavior, CS leverages existing knowledge on what constitutes high quality professional development. High quality professional development requires content focus, active learning opportunities, coherence, sufficient duration, and opportunities to discuss ideas with others teaching the same content (Desimone, 2009). Further, CS includes coaching since recent research points to the importance of coaching for improving fidelity of implementation (Reinke et al., 2014) and student behavioral outcomes (Stormont et al., 2015). Despite the impact of CS on CS practices, this study was conducted during teachers' first year of CS use. This is typical in intervention studies but means we know too little about how teachers show increases or decreases in their use of programs over time.

Impact of CS on student outcomes

Students in CS classrooms demonstrated more knowledge about energy, electricity, and renewable and nonrenewable energy sources compared to students in the "business-as-usual" group who received the same content through kits and standard curriculum. The science achievement effect exceeds the WWC threshold (0.25) for "a substantively important" effect (WWC, 2017, p. 14). A meta-analysis by J. Taylor

et al. (2018) examined the magnitude of intervention effects from 242 science intervention randomized-controlled trials, revealing a mean standardized difference of 0.41 for elementary school programs, slightly higher than 0.32 for CS. However, the magnitude of CS findings are comparable to other science-related PBL (e.g., Harris et al., 2015, reporting 0.25).

Findings showed that students participating in CS outperformed students in the control group on energy attitudes and behaviors. Energy attitudes and behaviors are highly valued outcomes of science SL because increased knowledge without shifts in attitudes and behaviors does not typically motivate students to take action and create change (Hollweg et al., 2011; Steg & Vlek, 2009). The impact of CS on energy attitudes and behaviors fits with efforts in science education to increase the relevance of science learning so that student notice problems and take action in the world (Roth and Calabrese Barton, 2004).

It is worth noting that findings showed statistically significant impacts on students' domain-specific civic engagement (energy attitudes and behaviors), but not domain-general civic engagement (civic skills and efficacy). Domain-specific civic engagement may emerge after a single SL experience but it may take numerous experiences as civic actors with opportunities to discuss successes and failures for students to develop a broad belief that they have the skills and capability to create impacts beyond the scope of their specific experiences. Future longitudinal work on SL can explore the development of civic engagement as students experience a variety of ways to work toward a better world (Sherrod, 2015).

Stronger findings for energy attitudes and behaviors versus civic skills and efficacy may stem from a challenge that occurs when you present students with authentic problems. Engaging with authentic problems means that students experience the reality that it is difficult to produce change. As a result, they may size up their skills and efficacy more conservatively. For example, pertaining to this data collection, one class chose a school-wide "electronic challenge" where all students at the school received a flyer on which they noted which electronic device their families would give up for a full week to save electricity. Although 37% of the families accepted the challenge, the negative comments from families (e.g., "we tried but it was too hard") were disheartening to students. In this study, students found themselves working on very challenging problems in SL. Going forward, teachers will need to help students realize small signs of progress to cultivate civic efficacy and prevent discouragement.

Use of Connect Science practices

Teachers in the intervention group used more CS practices than those in the control condition, indicating that the professional development, materials and coaching produced change in practice. Measuring actual use of CS practices (and CS-like practices in the control group) is an important step to confirm treatment differentiation between intervention and control groups. Results showed that CS led to use of CS practices, enabling us to use mediation analyses to consider the extent to which CS practices explained outcomes.

The findings from the mediation analyses suggested that actual use of CS practices explained the relation between CS and energy attitudes and behavior outcomes. Thus, engagement in an energy-related SL project cultivates energy attitudes and behaviors. By taking action toward solving a problem, students apply new concepts learned about energy to their everyday lives, which produces the affective and behavioral changes that reflect deep learning. This premise fits with meta-analytic work showing the importance of working with others to consolidate learning (Hattie & Donaghue, 2016). It also aligns with developmental work that draws attention to the socialization experiences that lead to early civic habits, and in turn, later civic engagement (Mitra & Serriere, 2012; Sherrod, 2015). These results stand in contrast to analyses showing that NGSS instructional practices did not mediate the relation between CS and energy attitude and behaviors. NGSS practices may produce science learning but fall short in producing the affective experience needed for students to believe that saving energy is important and to talk to their families about ways to save energy.

Use of CS practices mediated the relation between CS and energy attitudes and behaviors, but not science achievement. The lack of findings were surprising and require further inquiry. As with most new multi-component interventions, it is not clear which aspects of CS are core components of the intervention. Further work needs to be done to identify fundamental kernels of practice (Abry et al., 2014; Embry & Biglan, 2008; Jones et al., 2017). Also, to assess CS practice we asked teachers, not students, to report on classroom practices. These findings may reflect a discrepancy between what teachers report they are doing and students' actual uptake of those experiences.

Moderation findings for social skill outcomes

The lack of main effects of CS on communication and social competence was surprising. The finding that the social skill improvement was only evident in classrooms implementing with high fidelity suggests some association between use of CS and these student outcomes. The relation between CS dosage and social competence in the intervention group offers further support for this association. However, the correlational findings mean we are unable to make a causal claim. The lack of main effect may reflect the nature of universal interventions that are designed to teach skills to all students in a classroom, not only those at risk for problems. The effects of such programs are difficult to detect in small samples and in the short-term (Greenberg & Abenavoli, 2017). Pertaining to sample size, teachers reported on only six students per classroom for this outcome. To achieve 0.80 power, social skill effect sizes would need to exceed $d = 0.53$. Future work with larger sample sizes and long-term follow-up is needed.

Yet another explanation stems from the use of teacher report measures resulting in reference bias. When teachers report on student behavior, they use a frame of reference that may vary systematically between the intervention and control condition (Duckworth & Yeager, 2015). Because CS focuses on social skill instruction, intervention teachers may hold higher standards for their students and rate them more stringently. Future work could use vignette measures (Gasser et al., 2014) or direct assessment of SEL using computer-based assessment tools (McKown et al., 2016).

Limitations and future directions

The limitations of this study require mention. First, we were unable to collect baseline data on classroom practices, science achievement, civic engagement, and social skills to be included as covariates. As a result, the analyses were less precise at measuring treatment effects. Second, the variation in classroom schedules required that data collection windows be several weeks long. We carefully weighed the options of conducting measurement at fixed points in time (e.g., 13 weeks) versus linking measurement to benchmarks in a lesson sequence as teachers proceeded through the CS and energy units. We decided to link measurement to benchmarks but realize that this is not the perfect solution to this common problem in school-based research. Third, despite a rigorous randomization process, the intervention group had higher percentages of EL students than the control group. To address the difference, we included percentage of EL students in the classroom as a covariate but acknowledge that this does not fully address the difference. EL may be a proxy for cultural differences that were not fully unpacked in these data. It is also possible that EL students experience

marginalization, feel less connection to their community, and feel less agency to become civically engaged (Sherrod, 2015). Fourth, it was difficult to identify age-appropriate measures of the key constructs. We used systematic approaches to adapt and pilot measures, yet the final measures had lower than ideal internal reliability (alphas = 0.65–0.68). School-based research demands measures that have as few items as possible but still produce reliable, valid results; the balance between parsimony and reliability/validity is challenging to achieve (West, Buckley, Krachman & Bookman, 2018). The challenges were greater given the high percentage of English Learners in the sample. Future civic engagement research needs to contend with issues of measurement invariance across groups and consider alternative approaches to data collection in samples with a high number of ELs.

Existing work points to the importance of coherence between classroom-based programs and broader school efforts (NRC, 2012; Oberle et al., 2016; Whitworth & Chiu, 2015). Although beyond the scope of this study, we need more research attention on the district contexts and teacher attributes that forecast variation in implementation of CS. Not all districts prioritize instruction related to environmental issues. Teachers feel pressured to cover reading and math on state tests, which can lead to superficial coverage of lessons or difficulty implementing SL. Some teachers feel greater urgency to teach students about environmental crises and dwindling natural resources whereas others may lack knowledge or be afraid that these discussions will create anxiety or tension among students. These barriers raise questions about what supports need to be in place for all teachers to achieve high implementation of programs like CS.

Conclusion

This study offers useful insights about SL that integrates science and SEL. It extends the existing literature by focusing on SL in elementary school students (Celio et al., 2011) and by honing in on energy-related problems which can be challenging to teach because of the level of abstraction. Similar whole child approach programs have been investigated in descriptive and theory-generating research (Mitra & Serriere, 2012) and with pre- post-designs (Condon & Wichowsky, 2018). This study extends knowledge by using a research design that allows causal inference. Results show effect sizes that suggest almost a third of a standard deviation gains compared to business-as-usual. These effect sizes (0.32 for achievement, 0.31 for energy attitudes and behaviors) are above the What Works Clearinghouse standard for substantively important, roughly comparable to other science-based and project-based interventions, and suggest promise for schools interested in boosting science achievement while also enhancing students' development of the attitudes and behaviors that contribute to environmental consciousness and action. Findings suggest that when teachers give students a chance to work together and apply their new knowledge and skills to a problem that they care about, deep learning can result.

Acknowledgements

The research reported here was supported by the Institute of Education Sciences (IES), U.S. Department of Education, through grants R305A150272 and R305B140026 to the University of Virginia. The opinions expressed are those of the authors and do not represent views of the IES or the US Department of Education. We thank Tracy Harkins, Rebecca McGregor, Kristen Jones, Jacqueline Fisher, our graduate and undergraduate students, and the collaborating district administrators, teachers, and students.

Appendix A. Appendix: Connect Science Overview

Connect Science focuses on six essential understandings. Connect Science is based upon an 8-Step Framework which are listed here with a description of activities and instruction in each step. The lessons are more scripted and specific in the early steps of the sequence but emphasize teacher facilitation and offer more room for student voice in the later steps.

Six Essential Understandings	
1. Systems thinking can be useful in understanding interactions in the world and designing solutions to challenging problems. 2. Energy is present in different forms as it moves through natural and human-made systems. 3. Limited amounts of natural resources are available on earth. Each decision we make about our use of natural resources can have positive or negative impacts on the environment and other people. 4. Every citizen has a responsibility to find creative solutions to problems they notice in the world around them. 5. The best solutions arise when people with different knowledge and perspectives listen to each other, communicate respectfully, and collaborate to solve problems. 6. Kids can use their skills and knowledge to improve their community and our world by engaging in a service-learning project.	
Step 1: Define Service-Learning	
1.1 Establishing Classroom Norms 1.2 Citizens Who Make a Difference 1.3 Active Listening 1.4 Respectful Communication	Step 1 lessons introduce students to service-learning and create excitement about a future service-learning project. The lesson is designed so students begin to see how they can use their knowledge and skills to take action in their community. These lessons offer guidance to establish classroom norms and teach essential collaborative skills
Step 2: Discover Needs and Problems	
2.1 What do you Know about Energy and Resources? 2.2 Exploring Energy Systems 2.3 Building Circuits 2.4 Energy in Power Plant Systems 2.5 Energy from Fossil Fuels 2.6 Respecting Multiple Perspectives 2.7 Renewable vs. Non-Renewable Resources 2.8 Energy for the Future	In Step 2, students learn how to make connections between natural resources and the energy they use in their daily lives. Students explore systems and learn to recognize energy in different forms in the world around them.
Step 3: Investigate Problems	

(continued on next page)

(*continued*)

3.1 Managing Strong Emotions 3.2 How Far Does Your Food Travel? 3.3 Classroom Energy Audit 3.4 Energy Use in Our Daily Lives 3.5 Giving and Receiving Feedback 3.6 Energy and Resources RAFT Assessment 3.7 Focusing Student Investigations	In Step 3, students investigate energy and resource use problems. Students learn how their energy use can impact the environment. These lessons help students develop interest and urgency to address problems related to energy use. This step includes a RAFT (Role, Audience, Format, Topic; Tomlinson, 2003) assessment allowing the teacher to understand student understanding. This step also prepares students to manage strong emotions, which is an important skill for them to develop before working with each other on their service-learning project.
Step 4: Research Solutions	
4.1 Multiple Approaches to Solving Problems	In Step 4, students research different ways to have impact in their community. This step teaches three common approaches to creating change: 1) working directly on the problem, 2) educating others about a problem, or 3) working with leaders in the school or community to change policy to address the problem.
Step 5: Decide on a Project	
5.1 Making Informed Decisions 5.2 Data Collection Decisions 5.3 Perseverance	Step 5 lessons help students choose a solution and decide on a feasible project that will have a positive impact on their chosen energy problem. One lesson provides guidance on how to collect data to determine project impact. The other lesson helps students develop skills to persevere, even when the work gets difficult.
Step 6: Plan the Project	
6.0 Creating an Action Plan Series Summary 6.1 Identifying Tasks and Seeking Resources 6.2 Understanding and Sharing our Project Description 6.3 Creating a Timeline and Designating Roles 6.4 Being Aware of Progress	Step 6 is designed to help students develop essential project management skills. Each lesson teaches a component of developing an action plan. Students write a project description, which serves as a formative assessment. This step also ensures that students are actually learning what the project is intended to teach. These lessons offer strategies for identifying tasks and resources and distributing roles and responsibilities.
Step 7: Implement the Plan	
7.1 Resolving Conflict 7.2 Evaluating our Impact	In Step 7, students work together to implement their service-learning project. This step may take two or more weeks because this is the time when students are actually engaged in project work. One lesson offers a 4-Step strategy to identify and resolve conflict. The other lesson offers ways to help teachers and students monitor and reflect on their progress.
Step 8: Evaluate Impacts	
8.1 Our Class of Citizens who Make a Difference	Step 8 marks the end of students' service-learning experience. This step includes a lesson for students to evaluate the impact of their project on the original energy problem. The teacher and students reflect on their service-learning experiences and envision ways they can be engaged citizens in the future.

References

Abry, T., Hulleman, C. S., & Rimm-Kaufman, S. E. (2014). Using indices of fidelity to intervention core components to identify program active ingredients. *American Journal of Evaluation, 36*(3), 320–338.

Amna, E. (2012). How is civic engagement developed over time? Emerging answers from a multidisciplinary field. *Journal of Adolescence, 35*, 611–627.

Baroody, A. E. (2017). Exploring the contribution of classroom formats on teaching effectiveness and achievement in upper elementary classrooms. *School Effectiveness and School Improvement, 28*(2), 314–335. https://doi.org/10.1080/09243453.2017.1298629

Billig, S., Root, S., & Jesse, D. (2005). The relationship between quality indicators of service-learning and student outcomes: Testing professional wisdom. In S. Root, J. Callahan, & S. Billig (Eds.), *Advances in service-learning research: Vol. 5. Improving service-learning practice: Research on models to enhance impacts* (pp. 97–115). Greenwich, CT: Information Age.

Bodzin, A. M., Fu, O., Peffer, T. E., & Kulo, V. (2013). Developing energy literacy in US middle-level students using the geospatial curriculum approach. *International Journal of Science, 35*(9), 1561–1589.

Calabrese Barton, A. C., Tan, E., & Rivet, A. (2008). Creating hybrid spaces for engaging school science among urban middle school girls. *American Educational Research Journal, 45*(1), 68–103.

Caswell, L., Billig, S., Goodson, B., Gan, K., Levin, M., & Unlu, F. (2011). *Toolkit for the Evaluation of Service Learning Programs. Prepared for the Corporation for National and Community Service*. Cambridge, MA: Abt Associates Inc.

Celio, C. I., Durlak, J., & Dymnicki, A. (2011). A meta-analysis of the impact of service-learning on students. *Journal of Experiential Education, 34*(2), 164–181.

Center for Youth and Communities. (2011). Evaluation of Youth as Citizen Scientists (YACS). In *A Learn and Serve America Funded Project*. Waltham, MA: Brandeis University.

Century, J., & Cassata, A. (2016). Implementation research: Finding common ground on what, how, why, where, and who. *Review of Research in Education, 40*(1), 169–215.

Century, J., Rudnick, M., & Freeman, C. (2010). A framework for measuring fidelity of implementation: A foundation for shared language and accumulation of knowledge. *American Journal of Evaluation, 31*(2), 199–218.

Collaborative for Academic, Social and Emotional Learning. (2017). Core SEL Competencies. In *Retrieved from*. https://casel.org/core-competencies/.

Colley, C., & Windschitl, M. (2016). Rigor in elementary science students' discourse: The role of responsiveness and supportive conditions for talk. *Science Education, 100*(6), 1009–1038.

Condliffe, B., Quint, J., Visher, M. G., Bangser, M. R., Drohojowska, S., Saco, L., & Nelson, E. (2017). Project-based learning: A literature review. Washington. In *DC: MDRC*.

Condon, M., & Wichowsky, A. (2018). Effects of an inquiry-based science curriculum on STEM and civic engagement. *The Elementary School Journal, 119*(2), 196–222.

Conway, J. M., Amel, E. L., & Gerwien, D. P. (2009). Teaching and learning in the social context: A meta-analysis of service learning's effects on academic, personal, social, and citizenship outcomes. *Teaching of Psychology, 36*, 233–245.

Dane, A., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review, 18*(1), 23–45.

Davis, E. A., Palincsar, A. S., Smith, P. S., Arias, A. M., & Kademian, S. M. (2017). Educative curriculum materials: Uptake, impact, and implications for research and design. *Educational Researcher, 46*(6), 293–304.

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality, 19*(2), 109–134.

Deci, E. L., Vallerand, R. J., Pelletier, L. G., & Ryan, R. M. (1991). Motivation and education: The self-determination perspective. *Educational Psychologist, 26*(3–4), 325–346.

Denton, P., & Kriete, R. (2000). The First Six Weeks of School. In *Strategies for Teachers Series*. Greenfield, MA: Northeast Foundation for Children.

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

DeWaters, J. E., & Powers, S. E. (2011). Energy literacy of secondary students in New York State: A measure of knowledge, affect, and behavior. *Energy Policy, 39*(3), 1699–1710.

Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher, 44*(4), 237–251.

Embry, D. D., & Biglan, A. (2008). Evidence-based kernels: Fundamental units of behavioral influence. *Clinical Child and Family Psychology Review, 11*(3), 75–113.

Furco, A., & Root, S. (2010). Research demonstrates the value of service learning. *Phi Delta Kappan, 91*(5), 16–20. https://doi.org/10.1177/003172171009100504

Gasser, L., Malti, T., & Buholzer, A. (2014). Swiss children's moral and psychological judgments about inclusion and exclusion of children with disabilities. *Child Development, 85*(2), 532–548.

Goncy, E. A., Sutherland, K. S., Farrell, A. D., Sullivan, T. N., & Doyle, S. T. (2015). Measuring teacher implementation in delivery of a bullying prevention program: The impact of instructional and procedural adherence and competence on student responsiveness. *Prevention Science, 16*, 440–450.

Greenberg, M. T., & Abenavoli, R. (2017). Universal interventions: Fully exploring their impacts and potential to produce population-level impacts. *Journal of Research on Educational Effectiveness, 10*(1), 40–67.
Gresham, F., & Elliott, S. N. (2008). *Social skills improvement system (SSIS) rating scales*. Bloomington, MN: Pearson Assessments.
Hamilton, L. S., Doss, C. J., & Steiner, E. D. (2019). Support for social and emotional learning is widespread. Creative Commons Attribution 4.0 International Public License, RB-10064-BMGF. Retrieved from https://www.rand.org/pubs/research_briefs/RB10064.html.
Harkins, T., Merritt, E., Rimm-Kaufman, S. E., Hunt, A., & Bowers, N. (2019). *Connect Science. Unpublished Manual, University of Virginia, Arizona State University, & Harkins Consulting*. LLC. Virginia: Charlottesville.
Harris, C. J., & Rooks, D. L. (2010). Managing inquiry-based science: Challenges in enacting complex science instruction in elementary and middle school classrooms. *Journal of Science Teacher Education, 21*(2), 227–240.
Harris, C. J., Penuel, W. R., D'Angelo, C. M., Haydel Debarger, A., Gallagher, L. P., Kennedy, C. A., … Krajcik, J. S. (2015). Impact of project-based curriculum materials on student learning in science: Results of a randomized controlled trial. *Journal of Research in Science Teaching, 52*(10), 1362–1385.
Hattie, J. A., & Donoghue, G. M. (2016). Learning strategies: A synthesis and conceptual model. *Science of Learning, 1*, 1–13.
Hayes, K. N., Lee, C. S., DiStefano, R., O'Connor, D., & Seitz, J. C. (2016). Measuring science instructional practice: A survey tool for the age of NGSS. *Journal of Science Teacher Education, 27*(2), 137–164.
Hollweg, K. S., Taylor, J. R., Bybee, R. W., Marcinkowski, T. J., McBeth, W. C., & Zoido, P. (2011). *Developing a framework for assessing environmental literacy*. Washington, D.C: North American Association for Environmental Education.
Hopkins, H. J. (2013). *The Tree Lady*. San Diego, CA: Beach Lane Books.
Huang, F. L., & Moon, T. R. (2009). Is experience the best teacher? A multilevel analysis of teacher characteristics and student achievement in low performing schools. *Educational Assessment, Evaluation and Accountability, 21*(3), 209–234.
Hulleman, C. S., & Cordray, D. S. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness, 2*(1), 88–110.
Hunt, A., Rimm-Kaufman, S. E., Merritt, E. G., & Bowers, N. (2020). Because the sun is really not that big. An exploration of fourth graders tasked with arguing from evidence. *Elementary School Journal, 121*(2), 256–282.
Jones, S., Bailey, R., Brush, K., & Kahn, J. (2017). Kernels of practice for SEL: Low-cost. In *low-burden strategies*. New York, NY: The Wallace Foundation.
KIDS Consortium. (2011). KIDS as Planners: A guide to strengthening students, schools and communities through service-learning. In *Revised and Expanded* (3rd ed.). Harkins Consulting, LLC: Lewiston, ME.
KIDS Consortium & Center for Youth and Communities. (2008). *Service-Learning Classroom Activity Report*. Unpublished measure.
Kokotsaki, D., Menzies, V., & Wiggins, A. (2016). Project-based learning: A review of the literature. *Improving Schools, 19*(3), 267–277.
Lieberman, G., & Seydel, J. (2019). P4BL: Linking phenomena-based, place-based, project-based, and problem-based pedagogy to deepen learning. In *Retrieved from*. https://catalyst.greenschoolsnationalnetwork.org/gscatalyst/june_2019/MobilePagedReplica.action?pm=1&folio=35#pg35.
Linn, M. C., Gerard, L., Matuk, C., & McElhaney, K. W. (2016). Science education: From separation to integration. *Review of Research in Education, 40*(1), 529–587.
Manz, E. (2015). Representing student argumentation as functionally emergent from scientific activity. *Review of Educational Research, 85*(4), 553–590.
McKown, C., Russo-Ponsaran, N. M., Johnson, J. K., Russo, J., & Allen, A. (2016). Web-based assessment of children's social-emotional comprehension. *Journal of Psychoeducational Assessment, 34*(4), 322–338. https://doi.org/10.1177/0734282915604564
McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods, 22*, 114–140.
Mihaly, K., & Mccaffrey, D. F. (2015). Grade-Level variation in observational measures of teacher effectiveness. In K. A. K. Kane, & R. C. Pianta (Eds.), *Designing teacher evaluation systems: New guidance from the Measures of Effective Teaching Project* (pp. 7–49).
Mitra, D. L., & Serriere, S. C. (2012). Student voice in elementary school reform: Examining youth development in fifth graders. *American Educational Research Journal, 49*(4), 743–774.
Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2016). Science achievement gaps begin very early, persist, and are largely explained by modifiable factors. *Educational Researcher, 45*(1), 18–35.
Morningside Center. (2019). The 4Rs Program. In *Retrieved from*. https://www.morningsidecenter.org/4rs-program.
Muthén, L. K., & Muthén, B. O. (1998–2010). *Mplus User's Guide. Sixth edition [computer software and manual]*. Los Angeles, CA: Author.
National Academies of Sciences, Engineering, and Medicine [NASEM]. (2018a). *How People Learn II: Learners, Contexts, and Cultures*. Washington, DC: The National Academies Press.
National Academies of Sciences, Engineering, and Medicine [NASEM]. (2018b). *English Learners in STEM Subjects: Transforming Classrooms, Schools, and Lives*. Washington, DC: The National Academies Press. https://doi.org/10.17226/25182
National Assessment of Educational Progress. (2015). Retrieved from https://www.nationsreportcard.gov/science_2015/#?grade=4.
National Research Council. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: The National Academies Press. https://doi.org/10.17226/13165
National Youth Leadership Council. (2008). K-12 Service Learning Standards. In *Retrieved from*. https://www.nylc.org/page/standards.
Oberle, E., Domitrovich, C. E., Meyers, D. C., & Weissberg, R. P. (2016). Establishing systemic social and emotional learning approaches in schools: A framework for schoolwide implementation. *Cambridge Journal of Education, 46*(3), 277–297.
Peugh, J. L., & Enders, C. K. (2004). Missing data in educational research: A review of reporting practices and suggestions for improvement. *Review of Educational Research, 74*(4), 525–556.
Pruitt, S. L. (2014). The Next Generation Science Standards: The features and challenges. *Journal of Science Teacher Education, 25*(2), 145–156.
Reinke, W. M., Stormont, M., Herman, K. C., & Newcomer, L. (2014). Using coaching to support teacher implementation of classroom-based interventions. *Journal of Behavioral Education, 23*, 150–167.
Reiser, B. (September, 2013). *What professional development strategies are needed for successful implementation of the Next Generation Science Standards? Presentation at Invitational Research Symposium on Science Assessment*. Northwestern University. Retrieved from https://www.ets.org/Media/Research/pdf/reiser.pdf.
Rimm-Kaufman, S. E., & Merritt, E. G. (2019). Let's power our future: Integrating science and social and emotional learning to improve collaborative discourse and science understanding. *Science and Children, 57*(1), 52–60.
Rogoff, B. (1994). Developing understanding of the idea of communities of learners. *Mind, Culture, and Activity., 1*(4), 209–229.
Roseman, J. E., Herrmann-Abel, C. F., & Koppal, M. (2017). Designing for the next generation science standards: Educative curriculum materials and measures of teacher knowledge. *Journal of Science Teacher Education, 28*(1), 111–141.
Roth, W. M., & Calabrese Barton, A. (2004). Rethinking Scientific Literacy. In *Retrieved from*. https://ebookcentral-proquest-com.proxy01.its.virginia.edu.
Sanson, A. V., Van Hoorn, J., & Burke, S. E. L. (2019). Responding to the impacts of the climate crisis on children and youth. *Child Development Perspectives, 13*(4), 201–207.
Serrierre, S. C. (2014). The role of elementary teacher in fostering civic efficacy. The. *Social Studies, 105*, 45–56.
Sherrod, L. R. (2015). Recognizing civic engagement as a critical domain of human development. *Research in Human Development, 12*(3–4), 312–318. https://doi.org/10.1080/15427609.2015.1068040
Snyder, J., & Lit, I. (2010). Increasing the application of knowledge about child and adolescent development and learning in educator preparation programs. In *Paper commissioned by the National Expert Panel on Child and Adolescent Development and Learning in Educator Preparation of the National Council for Accreditation of Teacher Education (NCATE)*.
Stormont, M., Reinke, W. M., Newcomer, L., Marchese, D., & Lewis, C. (2015). Coaching teachers' use of social behavior interventions to improve children's outcomes: A review of the literature. *Journal of Positive Behavior Interventions, 17*(2), 69–82.
Taylor, J. A., Kowalski, S. M., Polanin, J. R., Askinas, K., Stuhlsatz, M. A. M., Wilson, C. D., … J, S. (2018). Investigating science education effect sizes: Implications for power analyses and programmatic decisions. *AERA Open.*. https://doi.org/10.1177/2332858418791991
Tidd, S. T., Stoelinga, T. M., Bush-Richards, A. M., De Sena, D. L., & Dwyer, T. J. (2018). An intensification approach to double-block algebra: A pilot implementation of Intensified Algebra in a large urban school district. *The Journal of Educational Research, 111*(1), 95–107.
Tomlinson, C. A. (2003). Fulfilling the promise of the differentiated classroom: Strategies and tools for responsive teaching. In *Alexandria*. Virginia: ASCD.
Trends, C. (2014). Measuring elementary school students' social and emotional skills. In *Retrieved from*. https://www.childtrends.org/wp-content/uploads/2014/08/2014-37CombinedMeasuresApproachandTablepdf1.pdf.
Vieira, R. M., & Tenrerio-Vieira, C. (2016). Fostering scientific literacy and critical thinking in elementary science education. *International Journal of Science and Math Education, 14*(4), 659–680.
Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. In *Cambridge*. Massachusetts: President and Fellows of Harvard College.
What Works Clearinghouse. (2015). Attrition Standards; WWC Standards Brief. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_brief_attrition_080715.pdf.
What Works Clearinghouse. (2017). Procedures Handbook. In *Version 4.0*. https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_procedures_handbook_v4_draft.pdf.
Whitworth, B. A., & Chiu, J. L. (2015). Professional development and teacher change: The missing leadership link. *Journal of Science Teacher Education, 26*(2), 121–137.
Wong, H. K., Wong, R. T., & Seroyer, C. (2009). The first days of school: How to be an effective teacher. In *Harry K Wong Publication*.
Wood, C. (2007). *Yardsticks: Children in the Classroom, Ages* (pp. 4–14). Turner Falls, MA: Northeast Foundation for Children.