"BECAUSE THE SUN IS REALLY NOT THAT BIG"

An Exploration of Fourth Graders Tasked with Arguing from Evidence

ABSTRACT

Little is known about the integration of social-emotional learning (SEL) with science instruction. We used a sequential mixed-methods design to examine (1) how fourth graders use argumentation practices and social gestures in science class and (2) how argumentation practices and social gestures differ between intervention and comparison classrooms. Intervention classrooms implemented Connect Science. Fourteen student conversations in seven classrooms were coded for argumentation practices (i.e., claims, evidence, and questions) and social gestures (i.e., agreement, disagreement, assertive speech, and prosocial speech). Across all classrooms, science conversations were most productive when students used social gestures to support use of argumentation practices. Without social gestures, conversations were disconnected or highly assertive. Proportionally, Connect Science students discussed science content more and discussed logistics less than comparison students. Findings include recommendations for conditions (i.e., SEL instruction, science reference materials, and time) to enhance scientific discourse and argumentation in elementary school classrooms.

Ashley Hunt
Sara E. Rimm-Kaufman
UNIVERSITY OF VIRGINIA

Eileen G. Merritt
VIRGINIA TECH

Nicole Bowers
ARIZONA STATE UNIVERSITY

THE ELEMENTARY SCHOOL JOURNAL
Volume 121, Number 2. Published online November 5, 2020

> Learning the fine art of speaking with the possibility of being heard, and listening with the possibility of being changed, is a practical contribution to finding one's way in a wildly diverse democracy.
>
> —Bill Ayers, *I Shall Create!*

MORE and more, society expects teachers to guide students' social and emotional development in addition to their academic learning. Nevertheless, students often progress through the public education system without acquiring critical social skills that underlie personal and professional success (National Research Council [NRC], 2012a). The social-emotional learning (SEL) research literature and practice guidelines identify effective approaches to supporting students' awareness of and ability to manage their attention, thoughts, and behaviors in school and beyond (Collaborative for Academic, Social, and Emotional Learning [CASEL], 2012; Durlak et al., 2011; Jones et al., 2017; Taylor et al., 2017). Several decades of research have established students who develop SEL skills are also more likely to flourish academically (Durlak et al., 2011; Taylor et al., 2017).

A growing body of work shows the potential for SEL curricula to teach social and emotional skills that transfer to countless settings throughout students' lives. Programs with evidence of effectiveness like PATHS (Promoting Alternative Thinking Strategies) and Positive Action provide teachers a structured curriculum of SEL lessons and activities addressing explicit social and emotional skills, behaviors, and attitudes (Kam et al., 2003; Lewis et al., 2016). Teachers lead activities like helping students recognize and label emotions or role-playing interpersonal situations (Jones et al., 2017; Rimm-Kaufman & Hulleman, 2015). However, the most clearly stated mission for schools remains focused on academic performance.

New work explores approaches that integrate SEL practices with academic instruction (Harris et al., 2015; NRC, 2012a). The links between SEL and science instruction in particular have noteworthy implications for practice. The Next Generation Science Standards (NGSS) position science as a discovery process through which students engage in scientific and engineering practices (SEPs) as they investigate phenomena (NRC, 2012b). Although typical teacher-centered science instruction has focused on memorization of vocabulary and following rigid procedures (Banilower et al., 2018; Reiser, 2013; Trygstad et al., 2013), NGSS describes a student-centered approach where students learn by engaging with scientific practices (NRC, 2012b). For example, scientific argumentation requires students to build collective knowledge about a topic supported by evidence through verbal or written arguments, implicating the use of social gestures like voicing agreement and disagreement (Hmelo-Silver et al., 2007). The SEPs described by NGSS place explicit demands on students' ability to articulate their understanding (or lack thereof) through discourse. Despite conceptual overlap and evidence that "dual-purpose" approaches to instruction predict improved academic, social, and emotional outcomes (Kochenderfer-Ladd & Ladd, 2016), science educators rarely integrate the goals of SEL.

Purpose of the Study

Young students need opportunities to build and practice the complex skills that allow them to respectfully engage in authentic science practices. At the same time, teachers need support to meet the ambitious learning goals of NGSS. The current study uses an exploratory sequential mixed-methods design to describe how small groups of fourth-grade students used argumentation practices and social gestures to facilitate rich science discussions. We also explore how conversations differed when teachers used Connect Science (CS), an NGSS-aligned curriculum that integrates SEL instruction, versus typical science instruction.

Review of Literature

Sociocultural learning theory asserts that a climate of openness, mutual respect, and normative communication helps students feel comfortable taking the social risks associated with making and evaluating scientific arguments (Colley & Windschitl, 2016; Vygotsky, 1978). Science discussions help students learn when group members verbally engage in a "shared thinking process" (Rogoff & Toma, 1997). Mortimer and Scott (2003) operationalized such a process as "dialogic-interactive" (D-I). Engaging multiple voices in talk about multiple ideas generates knowledge, whereas one-sided interactions about a single viewpoint do not. Accordingly, D-I discourse serves as the foundation for productive discussions.

NGSS Implementation

Typical science instruction in elementary classrooms often consists of vocabulary, memorization of facts, and other activities with limited conceptual depth. The NGSS call for the use of practices that promote inquiry, creative problem solving, and collaborative engagement with science content (NRC, 2012b). The standards provide a framework for creating educational experiences that mirror authentic practices of scientists through three dimensions: SEPs, disciplinary core ideas, and cross-cutting concepts. However, assessments and pacing guides have been slow to adapt to NGSS-aligned instruction (Trygstad et al., 2013). Standards have not been paired with necessary materials or professional support, an especially pressing problem in a time when the majority of the teaching force has not received training on inquiry approaches or three-dimensional instruction during their preservice education (Pasley et al., 2016). Despite the majority of states adopting standards inspired by NGSS since their release, science education in practice has, on average, changed little.

Although implementation combines the three dimensions of NGSS, our study of the teaching and learning of argumentation focuses on a subset of SEPs that, taken together, distinguish communication in science as a necessary aspect of learning in science:

1. Asking questions and defining problems (SEP 1)
2. Obtaining, evaluating, and communicating information from scientific texts (SEP 8)
3. Constructing explanations and designing solutions to explain phenomena or solve problems (SEP 6)
4. Engaging in argument from evidence (SEP 7)

Science instruction based on these practices requires students to demonstrate their understanding through discussion, which necessitates a higher level of language and communication skills (Lee et al., 2014). Teachers report insufficient time for quality implementation and feeling "overwhelmed and intimidated" by NGSS (Hanuscin & Zangori, 2016, p. 809; Penfield & Lee, 2010). Although resources to support implementation have become increasingly available (e.g., NRC, 2015), they have focused on describing experiences that promote intended learning outcomes; they do not provide instructional guidance on how to boost students' SEL skills to make NGSS come alive in the classroom. Leveraging knowledge of effective SEL programs and practices might help teachers become comfortable with more time- and language-intensive approaches to science teaching and learning.

Argumentation in Elementary Science

Scientists develop ideas through collaboration and critique in writing and through dialogue. The ability to discuss and improve scholarly work, often referred to as "talking science" (Lemke, 1998, p. 91), is itself a critical skill for scientists beyond their content expertise. Argumentation is a specific type of science discourse during which individuals demonstrate content understanding through the use of evidence, explanation, and reasoning (Erduran & Jimenez-Aleixandre, 2008). From a sociocultural perspective, scientific argumentation provides participants a conversational framework for establishing and evaluating ideas as a group (Tippett, 2009).

One way to describe effective scientific discourse uses a framework described by Mortimer and Scott (2003). Students' interactions in the classroom are classified across two dimensions: (1) dialogic, representing multiple ideas or perspectives, or (2) authoritative, representing a single point of view. They are also interactive and engage multiple voices or noninteractive and are characterized by a single voice. Accordingly, the framework describes four communicative approaches:

1. dialogic-interactive (D-I), or multiple ideas presented by multiple voices
2. authoritative-interactive (A-I), or one idea presented by multiple voices
3. dialogic-noninteractive (D-N), or multiple ideas presented by a single voice (e.g., a student makes and evaluates a claim without response or interruption)
4. authoritative-noninteractive (A-N), or one idea presented by a single voice (e.g., a student makes a claim that receives no response)

Interactions categorized as A-I are the most common in traditional instruction and discourse in the classroom: one voice (usually the teacher's) imparts facts and knowledge to students (Tippett, 2009). For example, a specific type of A-I interaction known as the Inquire-Response-Evaluate pattern has been observed commonly in classrooms (Cazden, 2001). Conversely, student-centered science instruction prompts predominantly D-I conversations where students learn by and through conversations with each other and their teacher (Driver et al., 2000; Manz, 2015).

The educational value of conversation goes beyond sharing facts; it also reflects a group's ability to reach scientific understanding. Over time and with practice, students can begin to use argumentation practices and social gestures to communicate and refine their understanding of science content (Berland, 2011; Duschl & Osborne, 2002). Although student-centered instruction might be more effective for engaging students in deep science learning than typical instructional approaches, guiding the development of skills that underlie productive discussions places unique demands on educators (Driver et al., 2000; Hayes & Trexler, 2016; NRC, 2012a). Because the ability to make and evaluate arguments using evidence implicates both communication skills and scientific knowledge, argumentation represents an opportunity to explore the integration of science and SEL instruction.

SEL

SEL refers to a growing body of work that identifies effective ways of developing students' awareness of and ability to manage their attention, thoughts, and behaviors in school. The CASEL developed one of the most widely used frameworks to characterize social-emotional skills as comprising five key competencies: self-awareness, self-management, social awareness, relationship skills, and responsible decision making (CASEL, 2017). Importantly, the abilities to learn and apply the skills, knowledge, and attitudes associated with SEL have intrinsic value as stand-alone competencies (Jones et al., 2017). In addition to developmental benefits, research has generated substantial evidence that implementation of programs and approaches to school-level SEL relate to a variety of positive academic and performance outcomes for students.

Summarizing decades of evaluation research has generated substantial evidence that school-level implementation of SEL programs and approaches relates to a variety of positive academic and social outcomes. On average, students in intervention schools outperformed comparison peers academically by 11 percentile points and were more likely to exhibit positive social behaviors such as interpersonal problem solving and perspective taking (Durlak et al., 2011). A second meta-analysis of later follow-up (6 months to 18 years later) found that SEL interventions had long-term impacts on a variety of outcomes, including positive peer relationships and less involvement with the justice system (Taylor et al., 2017).

Teachers influence students' social and emotional experiences in the classroom through their instruction, relationships with students, and implicit learning that happens when students witness how teachers manage their own thoughts and feelings (Schonert-Reichl, 2017). Jones et al. (2017) described the importance of teachers helping students practice basic SEL skills in a variety of contexts. For example, teachers might introduce sentence stems for students to use when agreeing and disagreeing as a first step toward students agreeing or disagreeing with peer claims related to science content. This bridging might be critical in middle childhood, when students prepare for the transition to middle school and adolescence, both of which are associated with increasingly complex social and emotional experiences (Eccles & Roeser, 2011).

About the Intervention

The current level of knowledge points to a need to bring together science and SEL instruction. The increasing number of states adopting science standards informed by NGSS has increased the urgency of developing evidence-based programming for teachers tasked with implementation. In response, our research team developed a new project-based learning curriculum called Connect Science as part of an Institute of Education Sciences Innovation and Development grant. The CS professional development (PD) experience prepares teachers to embed explicit instruction on social and emotional skills with science. The science lessons align with NGSS disciplinary core ideas in earth and physical science (PS3 and ESS3) and engage students in using NGSS SEPs.

The curriculum development project culminated in a randomized controlled trial (RCT), providing a unique opportunity to examine whether CS supported students' use of argumentation skills. In the current study, both intervention and comparison classrooms implemented lessons and activities designed to engage students in scientific discourse. Students in CS classrooms experienced a sequence of foundational SEL and science lessons prior to the observed discussions, allowing for comparison of discussions across the two groups.

Research Questions

Our exploration of small group science discussions in fourth-grade classrooms addressed two research questions: (1) Across both groups, how did students use argumentation practices and social gestures during small group discussions in science? and (2) How did discussions in CS classrooms differ from those in comparison classrooms? By first analyzing how students engaged in scientific argumentation across a variety of contexts, the study provides insight into the nature of rich discussions that bring young students into the sociocultural community of scientists. After describing the range of conversations, we investigate differences between the groups to identify if and when students leveraged social and emotional skills toward authentic science learning.

Method

Using an exploratory sequential mixed-methods design, we described and coded transcripts of 14 conversations between 2 and 5 students in 7 fourth-grade classrooms (2 conversations in each classroom; 4 classrooms in the intervention group). We began with qualitative analysis of conversational elements and themes. Next, we generated a data set of frequency codes quantifying students' use of argumentation practices, social gestures, and communicative approaches. Findings from analytic memos and qualitative analysis were compared against descriptive statistics. Finally, we explored systematic differences in student conversations between the intervention and comparison groups.

Participants

Recruitment. The research team partnered with a large, urban school district in the southeastern United States transitioning to NGSS to evaluate CS. Administrators and researchers recruited fourth-grade science teachers to participate via email. There were 32 teachers randomly assigned at the school level to either attend PD and implement CS during the 2017–2018 school year ($n = 18$) or to teach their typical science instruction using district-provided kits to address the same set of standards ($n = 14$). The research team conducted observations in each classroom, and teachers and students completed surveys about their experiences in science. Teachers received program training and materials at no cost in addition to a stipend for participating in the research. Teachers in the waitlist comparison group were invited to attend the PD program in summer 2018 and implement the curriculum during the 2018–2019 school year with developer support.

Sample selection. We used a purposeful sampling method that maximized the amount of possible data to explore while leveraging the group equivalence established by random assignment. Three criteria were applied to the full sample in the CS RCT: (1) availability of observation data, (2) clear audio of two small group discussions, and (3) fall 2017 implementation of the energy science unit. Table 1 summarizes the demographic characteristics of the seven classrooms (four intervention and three comparison classrooms) that met these criteria.

A check for representativeness found that the teachers sampled did not differ from the full RCT sample by education, years of teaching experience, recent PD in SEL or science, gender, or ethnicity (all $ps > .1$). School and classroom data were used to compare the selected sample to the full sample on enrollment, class size, prior achievement, socioeconomic composition, and departmentalized instruction. Classrooms differed from the full RCT sample on three demographic variables. The seven sampled classrooms had fewer Black students, more Latinx students, and more students designated as English learners than those in the full sample (all $ps < .05$).

Table 1. Participant Characteristics

	Count/Mean (*SD*)		
	Total ($N = 7$)	Intervention—Ms. Corbett, Ms. Grace, Ms. Jones, Ms. Woodward ($n = 4$)	Comparison—Ms. Green, Ms. Hurst, Ms. Spencer ($n = 3$)
Schools:			
Enrollment	240 (68)	246 (70)	234 (81)
Prior fourth-grade science achievement	.26 (.22)	.19 (.08)	.36 (.33)
Percent economically disadvantaged	.51 (.21)	.57 (.11)	.43 (.31)
Teachers:			
Years of experience	7.43 (5.38)	8.50 (5.07)	6.00 (6.56)
Master's degree	5	3	2
Recent SEL professional development	5	3	2
Recent science professional development	4	3	1
Female	7	4	3
White	7	4	3
Classrooms:			
Class size	23 (1)	23 (1)	22 (2)
Departmentalized science instruction	5	3	2
Percent female	.47 (.10)	.48 (.11)	.45 (.10)
Percent English learners[a]	.40 (.28)	.42 (.25)	.39 (.37)
Percent White	.26 (.29)	.13 (.07)	.43 (.42)
Percent Black[a]	.18 (.14)	.24 (.17)	.10 (.03)
Percent Latinx[a]	.41 (.27)	.41 (.27)	.33 (.33)
Percent Asian	.07 (.08)	.05 (.08)	.09 (.08)

Note.—*SD* = standard deviations; SEL = social-emotional learning.

[a] Differs significantly from randomized controlled trial (RCT) sample mean ($p < .05$).

Description of CS

The CS program consists of 30 lessons that teach science, SEL skills, and civic engagement. The eight-step service-learning experience was modeled after the *Kids as Planners* (2011) framework. Lessons were sequenced so that students first learned and practiced communication skills and then used those skills to discuss complex ideas in science. First, teachers guide students in discovering energy and resource problems in their community. Students then progress to identifying potential solutions, and finally they engage in a service-learning project to implement a solution to their chosen problem. Teachers received 5 days of professional development, curriculum materials, and coaching throughout implementation.

Data Collection

Observation data were collected during a 4-week window during fall 2017. Classrooms were video and audio recorded during a science lesson and small group discussion task. On-site research assistants gathered data, which were then transmitted to the research team electronically.

Classroom context. Science teachers in CS classrooms led a structured sequence of SEL lessons prior to the observation. Students began by collaboratively generating norms for their classroom (CS Lesson 1.1). Teachers then led a discussion about what it looks and feels like to listen to others, and students practiced active listening by paraphrasing (CS Lesson 1.3). Next, students learned sentence stems for agreeing and disagreeing with others respectfully (CS Lesson 1.4). Later, students built on these basic skills by discussing the more complex process of showing respect for multiple perspectives. After reading a book about two friends who resolve a conflict that arises from their differing cultural experiences, students brainstormed questions they could ask to better understand those with differing beliefs and ideas (CS Lesson 2.6; Connect Science Team, 2017). Because the sequence intertwined SEL and science lessons, students could practice skills in low-stakes situations prior to using them during science learning. The curriculum also provided resources like student worksheets and sample anchor charts to reinforce SEL skills and concepts during science lessons.

Teachers in the comparison group used their existing curricula, which included materials and lesson guides from science kits for teaching about electric circuits and natural resources. Although two of the three teachers in the comparison group reported receiving SEL-focused PD within the last 3 years, they were not given guidance on how to integrate those SEL strategies with their science instruction.

Observation context. Intervention teachers were observed while enacting CS Lesson 2.8, "Energy for the Future," which engaged students in a discussion about tradeoffs between various energy sources. The lesson began by introducing a table of pros and cons associated with renewable (e.g., hydropower, solar power) and nonrenewable (e.g., coal, petroleum) energy sources. Next, students were asked to sort the list "from those we should use the least to those we should use the most in the future" (Connect Science Team, 2017, p. 125). Teachers asked students to explain their reasoning to a partner using information from the pros and cons table to support their claims. Students were encouraged to rearrange their cards if their perspectives changed as a result of their discussion and then to report final decisions to the group. In these classrooms, 8 conversations between 2 and 5 students yielded 50 minutes of audio ($M = 6.25$, $SD = 3.78$).

Comparison teachers were asked to identify a lesson that included science instruction and an activity that required students to use comparison (e.g., pros and cons) to talk about lesson content in small groups for the classroom observation. The content covered in these classes included structural and behavioral adaptations of plants and animals, movement and relationships between objects in space, and pros and cons of technology (see Results section for more description of these lessons and activities). In comparison classrooms, 6 conversations between 3 and 5 students yielded 60 minutes of audio ($M = 10.00$, $SD = 3.19$).

Quantitative data collection. School, teacher, and classroom demographics were provided by the district. Teachers reported on their years of experience, whether they were departmentalized instructors (i.e., only taught science), and other descriptive information.

Procedure

Observation data collection and transformation. Research assistants followed a standardized protocol for conducting classroom observations. The camera was placed so that the teacher and as many students as possible were visible for the duration of the lesson. The researcher moved the camera to focus on a single group of students and placed audio recorders with groups when they began their discussions. Recordings of sufficient quality to understand and differentiate between speakers were transcribed by Rev Audio & Transcription. Transcripts were compared with corresponding audio and (when available) video data to clarify speakers and improve precision. The final set of transcripts consisted of 14 conversations (2 per classroom), each between a group of 2 and 5 students. After preliminary analyses, codes were transformed into quantitative data reflecting the frequency with which students used different scientific argumentation practices, social gestures, and communicative approaches.

Description of coding approach. We used an approach to assessing the normative pragmatics of science conversations similar to that used by Nielsen (2012). The procedure for applying codes to the transcripts included four steps (described and defined in Table 2). The first step was to define the two units of analysis: turns and interactions. A *turn* was the smallest unit of analysis and defined as everything said by a single speaker until another participant spoke, at which point a new turn began. An *interaction* was defined as a cluster of thematically related turns.

Step 2 of the coding approach involved identifying claims. Claims were defined using Kuhn and colleagues' model of idea units in argumentation: that is, "any assertion made with justification" (Kuhn et al., 2013, p. 464). Claims were coded and then subcoded as justified if the speaker included a rationale during the same turn. For Step 3, turns were coded for content (i.e., science related or logistical), argumentation practices (i.e., evidence and questioning), and social gestures (i.e., agreement and disagreement). In Step 4, each interaction was coded by communicative approach: (1) dialogic-interactive, (2) authoritative-interactive, (3) dialogic-noninteractive, or (4) authoritative-noninteractive (Mortimer & Scott, 2003). Figure 1 shows excerpts from a conversation to illustrate how codes were applied.

Turn	Transcript		Turn Codes	Interaction Code
1	Camila:	You know at the gas station that they used? That's natural gas. It's for your car.	Science-related claim with justification	
2	Troy:	No, it's gasoline.	Science-related disagreement	
3	Camila:	It's natural gas.	Science-related disagreement	
4	Troy:	I disagree with you because it's not natural gas, it's gasoline.	Science-related disagreement supported by empirical evidence	Dialogic-Interactive
5	Camila:	Yeah, it's natural gas and they dig it from the ground.	Science-related disagreement with generalized evidence	
6	Dylan:	Well, natural gas is different from-	Science-related disagreement	
7	Troy:	From gasoline.	Science-related assertive speech	
8	Dylan:	Yeah.	Science-related agreement	
14	Troy:	I think we use petroleum too much.	Science-related claim (not justified)	Authoritative-Non-interactive
15	Dylan:	Natural gas is like fire or something.	Science-related claim (not justified)	
16	Troy:	It's easy to catch on fire.	Science-related agreement	
17	Dylan:	Yeah. Because you rub two sticks together to make natural gas.	Science-related agreement with generalized evidence	Dialogic-Interactive
18	Troy:	Well, I don't know about that.	Science-related disagreement	
19	Dylan:	Yeah, to make fire you rub two sticks together.	Science related disagreement with generalized evidence	
39	Dominic:	I haven't even said anything yet because of y'all always talking.	Logistical claim with justification	Authoritative-Interactive
40	Camila:	Say some things.	Logistical assertive speech	

Figure 1. Excerpts from a student conversation demonstrating how codes were applied to turns and interactions.

Analysis

Qualitative analysis. Coders wrote analytic memos to describe the argumentation and social skills observed during student conversations. The first phase of analysis aligned with Ryu and Lombardi (2015) deeming any verbal contribution to a discussion as evidence of engagement in social learning. As such, we considered all claims, evidence, and questioning in our analysis regardless of content relevance. Audio recordings were transcribed and coded. Then the research team engaged in an iterative process of identifying, confirming, and refining emergent themes.

Quantitative analysis. Qualitative codes reflecting the frequency of argumentation practices, social gestures, and communicative approach skills in student conversations were transformed into quantitative data, generating two analytic data sets (see Table 2). The first data set includes 806 turns describing the content, argumentation practices, and social gestures in each transcript ($M = 163.93$, $SD = 101.44$). The second data set includes 283 interactions ($M = 22.64$, $SD = 12.68$), each categorized by communicative approach. Descriptive analyses quantified the observed patterns in scientific discourse and argumentation across classrooms. Chi-squared analyses were used to compare use of argumentation practices and social gestures between the intervention and comparison groups.

Results

Teachers enacted a science lesson that included a student discussion component, which served as the backdrop for observation and analysis of two research questions:

Table 2. Coding Steps, Definitions, Examples, and Frequencies

Step, Code	Mutually Exclusive Subcode(s)	Definition	Example Quote(s)	Average Frequency per Conversation (*SD*)
1. Define units of analysis:				
Turn		Everything said by one individual until another speaks.	"You're wasting tons of paper." "We should do this one. Look, it says, 'reduces carbon dioxide when burned, releases another gas.' That's not good."	163.93 (101.44)
Interaction		Cluster of turns related to a single claim.	Rodrigo: "Rabbits have large ears so they can hear above danger." Luca: "The body?" Rodrigo: "Oh yeah, they use the body."	22.64 (12.68)
2. Identify claims:				
Making claims		Any assertion made by an individual.	"*So that means that hydropower is bad* because it could kill fish."	18.86 (10.65)
	Justified	Inclusion of rationale in the same turn as a claim.	"So that means that hydropower is bad *because it could kill fish.*"	4.86 (2.93)
	Category			
3. Code turns:				
Content	Science-related	Related to presented science topic.	"And coal and wind turbines. People use all of these."	56.36 (37.07)
	Logistics[a]	Related to materials or expectations for assignment.	"Oh, you were supposed to bring your highlighter?"	20.57 (30.03)
	Off-task	Unrelated to science content or the assigned task.	"Do you want to have a sleep over this weekend?"	5.36 (6.16)

	Argumentation Practices			
4. Code argumentation practices:				
Using evidence		Grounds for belief or disbelief of a claim.	"This is one of those things that cause pollution."	11.00 (11.01)
	Empirical	Evidence based on observation of phenomena.	"Okay, so the petroleum . . . it says here that it's often used."	7.14 (9.40)
	Generalized[a]	Evidence based on an undefined person or group.	"They're probably even going to forget even how to use bottom few."	2.14 (2.14)
	Personal	Evidence based on personal preference.	"I think that one's bad."	1.50 (1.65)
Asking questions		A turn that explicitly solicits a response.	"Chase, what do you think?"	12.43 (9.94)
	Social Gestures			
5. Code social gestures:				
Expressing agreement		Expression of a similar opinion.	"I agree with Luke."	7.57 (4.70)
Expressing disagreement		Expression of a difference of opinion.	"Well, I don't know about that."	7.21 (9.20)
Using prosocial speech[a]		Name use; compliments; please/thank you.	"Thank you so much. Thank you, Lisa." "That looks really good!"	4.43 (6.93)
Using assertive speech[a]		Interruption of another speaker; use of directives	Dylan: "Natural gas is like fire or something like that and . . ."Troy: "It's easy to catch on fire.""Listen to the bad things about it."	10.36 (8.82)
6. Code interactions:				
Communicative approach	Dialogic-interactive (D-I)	Multiple points of view presented by multiple voices.	Luke: "Danger. It just produces, kind of, danger." Marco: "Oh yeah, like pollution." Luke: "Well also you could get shocked."	12.21 (5.66)
	Authoritative-interactive (A-I)	One point of view presented by multiple voices	Mason: "I also think that wind could be good." Christian: "Yeah, wind could be good."	4.29 (3.52)
	Dialogic-noninteractive (D-N)	Multiple points of view presented by one voice.	"We're just supposed to highlight the ones that we're interested in? We can do that."	1.57 (1.74)
	Authoritative-noninteractive (A-N)	One point of view represented by one voice.	"Solar panel's the second least."	4.36 (3.50)

Note.—*SD* = standard deviation.

[a] Emergent code was developed from early analysis; all other codes were determined a priori.

(1) Across both groups, how did students use argumentation practices and social gestures during small group discussions in science? (2) How did discussions in CS classrooms differ from those in comparison classrooms? In Table 3, the column labeled "RQ1" summarizes discussions that occurred in all seven classrooms. The columns labeled "RQ2" compare results between intervention and comparison group classrooms.

The next section provides key details to help situate the results within the instructional context. Next we describe the NGSS argumentation practices, social gestures, and communicative approaches across all seven classrooms. Finally, we evaluate how findings differed between the intervention and comparison groups.

Discussion Tasks and Related Content across All Classrooms

We operationalized "productive science conversations" as those where students used argumentation practices and social gestures to have a conversation related to science content that engaged all group members. Student contributions were characterized as either (1) science related (renewable and nonrenewable resources, movement and relationships between objects in space, pros and cons of technology), (2) logistical (assigning roles, identifying materials, clarifying teacher expectations), or (3) off-task. Aspects of the tasks set the stage for different types of conversations, as seen in this excerpt from an intervention classroom:

> **Ms. Jones:** Open up your bag, take out your cards. You're going to look at this pro and con sheet. And you're going to talk with your partners about which [energy source] you think would be the best . . . that you guys think we should use for the future.

In Ms. Jones's classroom, 84% of turns were science related, 11% concerned logistics, and 5% were off-task. The discussion task described in the CS manual gave students one discussion question without a single "right" answer and access to reference materials. Ms. Spencer, one of the comparison teachers, gave her students a similar task:

> **Ms. Spencer:** We want to make sure that people are agreeing on our sort. Remember our two categories . . . behavioral and structural . . . You can use your Venn diagram as a tool to help if you would like.

In Ms. Spencer's classroom, 70% of turns were science related, 25% were about logistics, and 5% were off-task. Ms. Jones and Ms. Spencer both presented a clear question, provided reference materials, and emphasized the need for students to reconcile differences of opinion. In contrast, another comparison classroom gave the following assignment:

> **Ms. Hurst:** Your question was what kind of patterns did you see in space? . . . think about what you did today and use words and pictures to show me what you learned. Now, some of the words you might want to use—rotate, orbit, revolve . . . you have 10 minutes to do that.

Table 3. Descriptive Statistics of Coding Results across All Classrooms and between Groups

Code	Mutually Exclusive Subcode	Percent (*SD*) Total (RQ1)	Percent (*SD*) Intervention (RQ2)	Percent (*SD*) Comparison (RQ2)	χ^2
Turn content:					
	Science related	.55 (.50)	.70 (.46)	.45 (.50)	49.03***
	Logistics	.36 (.48)	.15 (.36)	.49 (.50)	92.27***
	Off-task	.09 (.29)	.15 (.35)	.06 (.23)	17.58***
		Argumentation Practices			
Making claims:					
	Justified	.26 (.44)	.25 (.44)	.26 (.44)	.05
	Science related	.52 (.50)	.69 (.46)	.40 (.49)	21.60***
	Logistics	.39 (.49)	.17 (.37)	.54 (.50)	37.21***
	Off-task	.09 (.29)	.14 (.35)	.06 (.23)	5.09*
Using evidence:					
	Empirical	.66 (.47)	.58 (.50)	.74 (.44)	4.31*
	Generalized	.20 (.40)	.30 (.46)	.11 (.32)	7.93**
	Personal	.14 (.35)	.12 (.34)	.15 (.35)	.27
	Science related	.82 (.38)	.86 (.35)	.79 (.41)	1.32
	Logistics	.10 (.30)	.04 (.20)	.15 (.36)	4.88*
	Off-task	.08 (.27)	.10 (.30)	.06 (.24)	.67
Asking questions:					
	Science related	.48 (.50)	.65 (.48)	.48 (.49)	10.92**
	Logistics	.41 (.49)	.19 (.40)	.53 (.50)	18.35***
	Off-task	.11 (.32)	.16 (.37)	.09 (.29)	2.03
		Social Gestures			
Expressing agreement:					
	Science related	.75 (.43)	.86 (.35)	.68 (.47)	4.82*
	Logistics	.17 (.38)	.02 (.14)	.30 (.46)	15.07***
	Off-task	.07 (.25)	.12 (.33)	.02 (.13)	4.47*
Expressing disagreement:					
	Science related	.79 (.41)	.94 (.24)	.72 (.45)	6.92**
	Logistics	.21 (.41)	.06 (.24)	.28 (.45)	6.92**
	Off-task	0	0	0	0
Using prosocial speech:					
	Science related	.36 (.48)	.71 (.49)	.31 (.47)	4.45*
	Logistics	.56 (.50)	.29 (.49)	.60 (.49)	2.50
	Off-task	.08 (.27)	0	.09 (.29)	.69
Using assertive speech:					
	Science related	.40 (.49)	.53 (.50)	.31 (.47)	6.69*
	Logistics	.50 (.50)	.29 (.46)	.63 (.48)	16.01***
	Off-task	.10 (.31)	.18 (.39)	.06 (.23)	5.87*
		Communicative Approaches			
Interaction:					
	Dialogic-interactive (D-I)	.54 (.50)	.60 (.49)	.50 (.50)	2.55
	Authoritative-interactive (A-I)	.19 (.39)	.16 (.37)	.20 (.40)	.66
	Dialogic-noninteractive (D-N)	.08 (.27)	.07 (.26)	.08 (.27)	.06
	Authoritative-noninteractive (A-N)	.19 (.40)	.16 (.37)	.21 (.41)	1.08

Note.—*SD* = standard deviation.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Nearly three quarters (74%) of the turns in Ms. Hurst's classroom were about logistics, whereas 20% of the turns were science related. Again, off-task turns were the least common (6%). Ms. Hurst's instructions provided students a broad discussion topic and required them to produce a physical product. She listed vocabulary that students might use verbally but did not include a written component. Assignments that asked a focused, well-articulated question and included reference materials set the stage for more content-focused discussions.

Use of Argumentation Practices and Social Gestures Observed across All Classrooms

Several patterns emerged in how students approached the assignment. We describe the range of ways in which students used argumentation practices and social gestures during their conversations based on the synthesis of qualitative themes and descriptive analyses.

Argumentation practices. Instances of students making and justifying claims, using evidence, and asking questions were analyzed to understand how students in all seven classrooms engaged in NGSS argumentation practices. Results are reported in the RQ1 column of the argumentation practices section of Table 3.

Making claims. A claim functions as the beginning of an argument, and students used them to introduce new ideas. Students made an average of 19 claims per conversation ($SD = 10.65$), but only 26% of claims included justification. Claims were often followed by another student responding directly, as seen in Ms. Grace's class below:

Ann: The one that's used the least is coal.
Erin: Coal, like the little rocks?

Student turns that initiated an interaction by presenting a new idea all met the behavioral definition of the NGSS practice of making a claim but varied in the extent to which they contributed to building scientific knowledge. On average, 52% ($SD = 0.50$) of claims were science related and provided evidence of students' content understanding. Claims prompted continued group engagement when they were brief and specific (e.g., "Coal, it helps us get energy, too"; "I think they're going to use hydropower most"; "Well, gasoline and oil are running out"). Long or abstract claims were often interrupted or ignored by other members of the group.

Teacher participation in discussions helped students produce more complex science-related claims, as seen in the interaction below about animal adaptations:

Ms. Spencer: What's something that for sure is going to go in structural?
Rodrigo: The fox.
Ms. Spencer: What about the fox is a structural adaptation?
Diana: He has thick fur, so he can camouflage.

Ms. Spencer's first question elicits a correct response from Rodrigo, confirming his understanding of the difference between structural and behavioral adaptations. Her follow-up question prompted the group to build on Rodrigo's claim. In response, Diana made a more sophisticated claim that included justification.

More than one third of claims (39%, $SD = 0.49$) were related to logistics. These claims helped groups determine how to complete the assignment. For example, in Ms. Green's class, Lily began an interaction by making ("I think Krista should go . . .") and justifying (" . . . she hasn't talked any") a logistical claim. Although unrelated to science, Lily made a valuable contribution that helped manage social participation in the discussion.

Fewer than 10% of claims were categorized as off-task (9%, $SD = 0.29$). Some off-task claims referenced the recording equipment in the classroom (e.g., "Hey, they're recording us") whereas others covered a wide range of topics (e.g., "I feel like I have a golden necklace, look at my necklace"; "I love science, it's the most funnest part of the year"). Turns like these were included in our analysis because they met the behavioral definition of the NGSS practice of making claims. However, compared with science-related and logistical turns, they were not nutritive to science learning. Such claims often led to brief off-task interactions or were ignored by other group members.

Using evidence. Students used three types of evidence: empirical, generalized, and personal (see the mutually exclusive subcodes under "Using evidence" in Table 3). More than half (66%, $SD = 0.47$) of students' evidence use was empirical. In intervention classrooms, students often used information from the pros and cons table as empirical evidence. For example, after Jessica suggested that one of the energy sources was "good," John responded, "No, it's not. Look, releases carbon dioxide when burned." John expressed his disagreement with Jessica's claim and used empirical evidence from the reference material to support his point of view.

Students also supported claims with generalized evidence of what some undefined set of people think or do ("Yeah, people don't use it that much"). Generalized evidence was the second most common (20%, $SD = 0.40$). The remaining 14% ($SD = 0.35$) of evidence was based on personal preferences ("I just like staring at them," referring to wind turbines).

Asking questions. Questions represent a bid to add another voice to an interaction. When students asked questions, they created opportunities to build group understanding. Questions often prompted students to supply evidence for a claim:

John: This one is the greatest.
Max: No, I mean how?
John: Because look, it just uses the sun's electrical power.

Students asked an average of 12 questions per discussion ($SD = 9.94$). Overall, conversations contained more questions about the science content than logistics, χ^2 (1,806) $= 8.94$, $p = .003$. Logistical questions were also useful, such as when Chase asked, "Can I talk?"

Social gestures. Expressions of agreement and disagreement, assertive speech, and prosocial speech were coded to demonstrate how students used social gestures across the 14 conversations. Results are reported in the RQ1 column of the social gestures section of Table 3. Assertive speech (turns that interrupted another speaker or included a directive) was, on average, the most common social gesture ($M = 10.36$, $SD = 8.82$). Next in prevalence was agreement ($M = 7.57$, $SD = 4.70$) followed by disagreement ($M = 7.21$, $SD = 9.20$). Turns that included a peer's name, gave a compliment, or included please or thank you were categorized as prosocial; prosocial speech was the least frequent social gesture ($M = 4.43$, $SD = 6.93$).

Agreement and disagreement. When students stated whether they agreed with a peer's claim, they offered their point of view to the group's collective understanding. On average, students expressed agreement and disagreement equally, $t(13) = 0.22$, $p = .83$. However, the two gestures prompted distinct responses. Agreement often led to brief interactions about a single idea, as seen in an example from Ms. Woodward's class:

Mason: I also think that wind could be good.
Christian: Yeah, wind could be good.

Although Christian makes a social contribution by agreeing, he echoes Mason's claim rather than building on the understanding of the group. Conversely, expressions of disagreement introduced a conflicting viewpoint for the group to resolve. The dialogue below illustrates how a disagreement about the size of the sun prompted further discussion:

Shane: Because the sun is not really that big.
Robert: The sun's pretty big. Like way bigger.
Shane: The sun is the same size, as probably like the earth.
Robert: No, it's bigger! . . . didn't you hear Ms. Hurst say it's bigger? It's way bigger.

After hearing Shane's claim that "the sun is not really that big," Robert used evidence to support his disagreement. Although his evidence may seem unsophisticated, he draws from his empirical observations of Ms. Hurst's instruction to justify his perspective.

Assertive speech. Students interrupted each other and used directives more than they used other social gestures. The use of assertive statements in conjunction with other social gestures helped students to manage participation in the discussion. For example, after two students in Ms. Green's class continually interrupted and spoke over each other, Amanda suggested, "If someone's talking, just let them talk." During the same conversation, Meredith responded to a complicated claim by saying, "Okay, hold on . . . we need to write this down."

Prosocial speech. Students' use of names, compliments, and please or thank you were infrequent relative to other social gestures with two notable exceptions. Students used prosocial speech 24 times during one of the conversations in Ms. Hurst's classroom. The majority (67%) of prosocial speech related to logistics:

Lisa: You're doing so well, Ruby and Giana. I wish I could be like you.
Giana: Thank you so much. Thank you, Lisa.

Other examples from this group include when Jennifer made an error and apologized: "I'm sorry, Ruby. I messed up" and her compliment: "The sun looks really good." Although prosocial speech made for a polite, friendly discussion, it did little to support science learning.

One of the conversations in Ms. Green's classroom included 15 uses of prosocial speech. In this group, prosocial speech was primarily science related (60%). Students addressed peers by name, such as when Amanda said: "I disagree with Paul because let's say there are owls or squirrels or birds that live in that tree . . . it's like if someone just came to your house when you weren't there, and knocked down your house." Later, Krista made a similar contribution: "I agree with Amanda, and I disagree with Paul." Students used prosocial speech to specify who they were addressing, allowing group members to acknowledge the viewpoints of others while making their own contributions.

Conversational profiles of social gesture use. Across the seven classrooms, distinct patterns emerged in the use of social gestures, resulting in three conversational profiles: balanced, disconnected, and highly assertive. Differences in the amount and type of social gestures used revealed the range in students' approaches to the social dimension of the task (see Fig. 2).

Balanced conversations included frequent use of a variety of social gestures and were longer than disconnected and highly assertive conversations, $t(12) = -3.13$, $p = .009$. Nearly two thirds of social gestures in these discussions were expressions of disagreement (31%, $SD = 0.09$) and assertive speech (31%, $SD = 0.11$). The remaining gestures were expressions of agreement (28%, $SD = 0.10$) and prosocial speech (10%, $SD = 0.06$).

Six of the 14 student discussions fit this profile. One example took place between four students in Ms. Corbett's class: Jordan, Brynn, Charlie, and Xander. Jordan began by making a claim that he built on throughout the discussion: "Well this is how it should be, but I don't think people are going to do that. They're probably going to use solar, nuclear . . . people are going to get so advanced and think they're getting smarter when they're really being stupider . . . I'd rely on old techniques the most." Brynn responded by voicing her disagreement ("I think they're going to use hydropower the most"). Jordan and Brynn engaged in a lively back-and-forth about

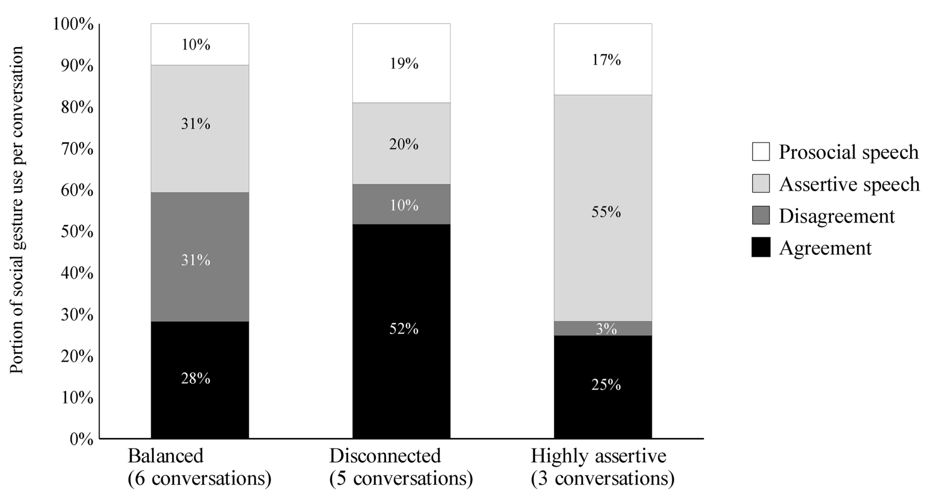

Figure 2. Composition of social gestures across all conversations representing three conversational profiles.

their views, supported by Xander, who made frequent social contributions despite not making his scientific point of view clear. When Charlie got the group off-task, Xander said: "We're supposed to be working together. You don't even have your cards out. Stop distracting us, we're trying to work." Here, Xander's assertiveness allowed the group to stay on task and continue their discussion about energy sources despite one member failing to contribute.

In contrast, disconnected conversations were brief and consisted of students using argumentation skills without social supports. Five of the 14 conversations fit this description. For example, consider the set of six interactions below comprising nearly one third (29%) of a brief discussion in Ms. Spencer's class:

Jared: Hawks have sharp claws that kill their prey.
Casey: What is this?
Molly: Bear?
Kiera: An artic fox has . . .
Molly: Insects are shaped like a leaf so predators think they are real leaves.
Jared: A rosebush has thorns to . . . where's this go?
Molly: Frogs have long strong legs to hop really far.

Even though all four group members participated, the discussion lacked evidence of students listening or responding to each other. Group members did not explicitly agree or disagree, and each contribution began a new, seemingly independent line of inquiry.

The remaining three conversations were highly assertive. These discussions were similar to disconnected conversations in length but were dominated by the presence of assertive speech (55%, $SD = 0.17$). Agreement was the next most common gesture (25%, $SD = 0.16$) but limited use of explicit expressions of disagreement (3%, $SD = 0.03$) did not balance the frequent assertive statements. As seen in Ms. Woodward's class:

Kamren: Guys, stop! We need to work.
Travis: He's not working, he didn't even highlight.
Ari: This all doesn't matter.
Kamren: We're supposed to be working, come on.
Travis: I already picked the most highest one.
Ari: Well, you chose the ugly thing bro.

Unlike when Xander used assertive speech to stop Charlie from derailing the discussion in Ms. Corbett's class, Kamren's attempts to keep the group on task and Travis's bid to redirect the conversation toward the science content were less effective in the absence of other social gestures. When the most prevalent social gestures used in a conversation were interruptions and use of directives, students did not have a productive dialogue.

Communicative Approaches across All Classrooms

Interactions were categorized as one of four communicative approaches: (1) dialogic-interactive, (2) authoritative-interactive, (3) dialogic-noninteractive, and (4) authoritative-noninteractive (Mortimer & Scott, 2003). Scientific argumentation prioritizes the

D-I communicative approach and, therefore, we expected that conversations with evidence of students using both argumentation practices and social gestures use would be primarily D-I. Results are reported in the RQ1 column of the communicative approaches section of Table 3.

Nearly three quarters of interactions (73%) engaged multiple voices. A chi-square test of independence revealed that within the portion of the interactions coded as interactive, more were dialogic than authoritative, χ^2 (1,338) = 72.81, $p < .001$. More than half (54%, $SD = 0.50$) of interactions included multiple voices expressing multiple points of view (D-I). Conversations in intervention classrooms, such as the following example from Ms. Grace's class, consisted of D-I interactions with connections to the lesson content:

Leslie: Every single day people use . . .
Erin: Gasoline to fill up their cars.
Sierra: Yeah, yeah, we might run out of it.

Regardless of content relevance, opportunities to practice D-I conversation can help students become comfortable with managing the type of dialogue required to build knowledge through scientific discourse (Berland, 2011). For example, the following interaction occurred in Ms. Hurst's class, where students were asked to "use words and pictures to show me what you learned":

Giana: You have to do the writing. I know how to do this.
Lisa: Here, I'll do the writing.
Jennifer: I'll do the outlining.

Here, different voices respectfully communicated their ideas about task roles. The students used D-I interactions to reach agreement on how to best complete the assignment rather than talking about what they learned about the relationships between objects in space.

Typical discourse patterns in elementary science prioritize multiple voices endorsing a single point of view (Tippett, 2009). Authoritative-interaction (A-I) communication was the second most common and accounted for nearly one quarter (22%, $SD = 0.39$) of interactions. One example of an A-I interaction took place in Ms. Green's class:

Luke: It's bad because people are digging up iron and copper for these utensils, which you can't make again. It's just gone, used for wiring. Just shorten that up.
Marco: Use too many resources.
Luke: Uses too many resources.

Although the interaction engaged two students, it did not represent an exchange of ideas, as Marco merely echoed Luke's claim.

Discussion Characteristics in Intervention versus Comparison Classrooms

The second research question compared differences between students' use of argumentation practices and social gestures among the intervention and comparison group classrooms. Results showed no difference in the frequency of argumentation practice or social gesture use between groups. However, the students in CS classrooms tended to have science-related conversations, whereas those in comparison classrooms talked more about logistics (see RQ2 column of argumentation practices and social gestures sections of Table 3). These comparisons are presented in Figure 3.

Students in intervention classrooms made more science-related claims (69%, $SD = 0.46$) than those in comparison classrooms (40%, $SD = 0.49$), χ^2 (1,806) = 21.60, $p < .001$. In comparison classrooms, more than half of claims (54%, $SD = 0.50$) were logistical, which was less than the 17% ($SD = 0.37$) in intervention classrooms, χ^2 (1,806) = 37.21, $p < .001$. Both groups used evidence to discuss science content, but students in comparison classrooms used evidence to support claims about logistics (15%, $SD = 0.36$) significantly more than the intervention group (4%, $SD = 0.20$), χ^2 (1,806) = 4.88, $p = .02$. Nearly two thirds of questions in intervention classrooms (65%, $SD = 0.48$) were about the science content, more than the 48% ($SD = 0.49$) in comparison classrooms, χ^2 (1,806) = 10.92, $p < .001$. Finally, students in comparison classrooms asked more logistical questions (53%, $SD = 0.50$) than those in intervention classrooms (19%, $SD = 0.40$), χ^2 (1,806) = 18.35, $p < .001$.

A similar pattern emerged in how the two groups used social gestures. Students in CS classrooms used all four of the observed social gestures to discuss science content more than students in comparison classrooms (all ps < .05). Conversely, students

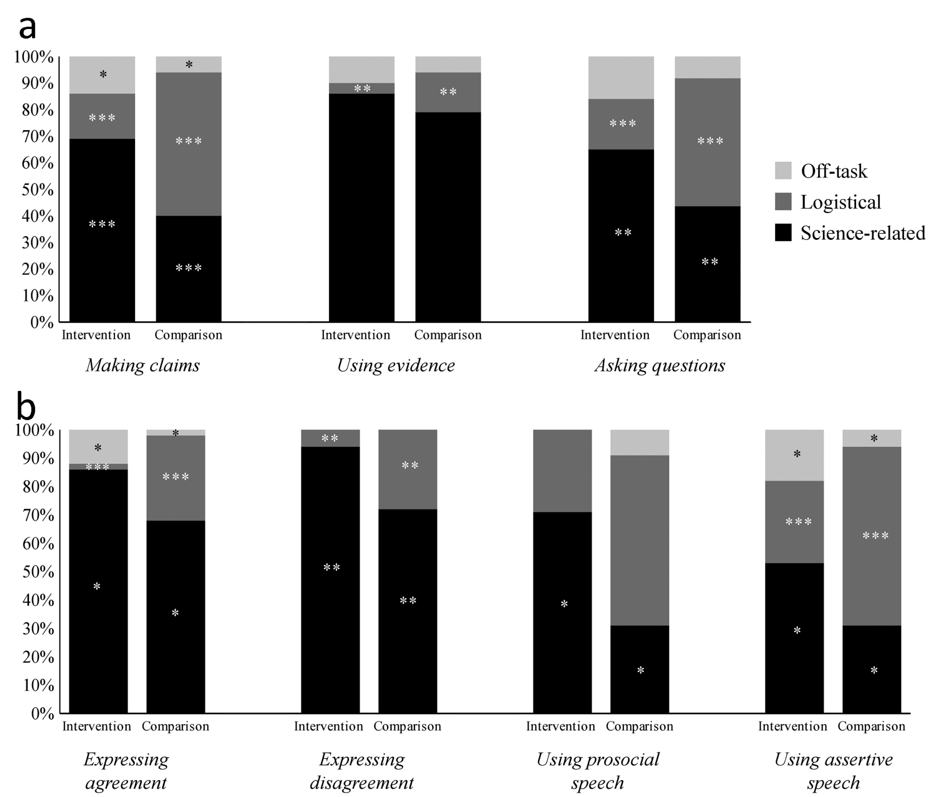

Figure 3. Differences in science-related, logistical, and off-task use of (*a*) argumentation practices and (*b*) social gestures between Connect Science and comparison classrooms.

in comparison classrooms expressed agreement, disagreement, and used assertive speech in discussion about logistics more than students in intervention classrooms (all $ps < .01$).

Discussion

Researchers can support elementary educators rising to the ambitious goals of NGSS by identifying ways to integrate SEL and science instruction. We described how 14 groups of 2–5 students in 7 fourth-grade classrooms used argumentation practices (e.g., making claims, using evidence, and asking questions) and social gestures (e.g., agreement, disagreement, assertive speech, and prosocial speech) during small group discussions in science. We also compared how students' use of argumentation practices and social gestures differed between two groups: four intervention classrooms implementing CS, an NGSS-aligned curriculum that integrates SEL, and three comparison classrooms implementing their typical science instruction. Our findings describe how students navigated the challenge of learning science through discourse. Furthermore, the group comparisons assisted in identifying conditions that encouraged productive science conversations among students in ways consistent with NGSS.

Common Discourse Practices across All Classrooms

The first research question sought to determine whether students were able to engage in productive science conversations, which we defined as discussions including an explicit focus on science content, use of argumentation practices and social gestures, and predominantly D-I conversations. The conversations that met these criteria demonstrate that fourth graders are capable of engaging in high-quality scientific discourse under certain conditions. For example, tasks must be designed with clear discussion outcomes in mind. Students need sufficient understanding of how content relates to their world, and they need to know how to use social gestures to sustain the dialogue. When these conditions were not met, conversations drifted toward the logistics of the assignment. When students did not use social gestures effectively, they had disconnected (i.e., two-way but unrelated to each other) or highly assertive conversations (i.e., one-way with students talking over each other). Subsequently, student conversations focused less on discussing and developing their ideas in ways that lead to deep-level thinking.

Claims, justification, and use of evidence. Scientific argumentation described by NGSS centers around the practice of making and justifying claims with evidence (NRC, 2012a). Though claims were common, only one in four was justified with evidence. Inclusion of reference materials seemed to push students toward supporting their claims with empirical evidence. For example, providing a handout with pros and cons of different energy sources prompted students to evaluate sources independently and objectively. Distributing a graphic organizer illustrating different animals and adaptations gave students specific examples to add to the discussion. Our findings extend the value of reference materials to an upper elementary context, as prior research with middle schoolers (Berland, 2011) and high schoolers (Nielsen, 2012) has demonstrated that reference materials encourage students to use evidence to support their claims.

Asking questions. Asking questions during discussions prompted group engagement. From a sociocultural perspective, questioning can function as a psychological tool that pushes students to think at a more sophisticated level (Hackling et al., 2010; Vygotsky, 1978). When students asked questions, they pressed peers to explain their thinking in a way that promoted deep-level learning. Those who used questioning effectively (i.e., to better understand a peer's claim) challenged others to unpack their claims and examine their reasoning. Questions were also inherently social, requiring that students listen to one another to build on the group's knowledge and understanding. Questioning also points to activation of a developmental process where students progress from one-sided conversation to dialogic conversations, moving them closer to the analytic conversations exemplified by scientists (Duschl & Osborne, 2002).

Using social gestures in science discussions. Groups that frequently used a variety of social gestures elevated their conversations, facilitated sustained discussion, and managed participation. Expressions of agreement and disagreement, assertive speech, and prosocial speech contributed to the rhythm of the discussion and allowed students to deepen their conversations about science. Several patterns emerged in how discussion groups used social gestures. Students interrupted each other and used directives across conversations, but these potentially delicate turns were productive when tempered by explicit expressions of agreement or disagreement and prosocial speech. When assertive speech dominated the conversation and students had limited time, the resulting conversations did little to build scientific knowledge. For example, a student in a highly assertive group might interrupt a peer to share a conflicting claim rather than making their claim after an explicit expression of disagreement.

The presence of prosocial language suggests a positive sense of community in the classroom that supports student learning (Jones et al., 2017). Although students' use of prosocial language indicated positive social development, prosocial speech was rarely on topic. The ideal balance between assertive and prosocial speech involved students using peers' names and giving compliments in a way that embedded the science content. For instance, one student responded to a peer's claim by saying, "Oh, I like that idea, that makes sense!" Productive science discussions involve a combination of prosocial and assertive speech such that the prosocial speech maintains a sense of community and connection while resolving disagreements that lead to collective science learning. When assertive speech was combined with other social gestures, students elevated the conversation to talk about scientific ideas without personalizing their position. We observed multiple cases of young students meeting this lofty goal. Nevertheless, groups that did not use social gestures effectively had less productive conversations. Students appeared to be talking at each other rather than having meaningful interactions.

Communicative approaches. Students' communicative approaches resembled the expectations for scientific argumentation described by NGSS. Interactions almost always engaged multiple voices, and individual bids for a response seldom went unanswered. D-I interactions (multiple points of view presented by multiple voices; Mortimer & Scott, 2003), which were the most useful for collectively building knowledge, were the most prevalent type of interaction. However, the next most common communicative approach was authoritative-interactive (A-I) wherein multiple voices discussed a single idea. This finding highlights the tendency for students to revert to authoritative discussion of facts and the need for continued support and scaffolding to promote primarily D-I conversations in science.

Findings from Group Comparisons

Teachers implementing CS set the stage for scientific discussion differently than the comparison teachers did. The groups did not differ in the quantity of argumentation practices and social gestures used. Instead, distinct patterns in how students used argumentation practices and social gestures emerged between the intervention and comparison groups.

Students in intervention classrooms used argumentation practices and social gestures to talk about science. As directed in the CS implementation manual, teachers gave a well-articulated discussion question and provided reference materials. In intervention classrooms where students appropriately used social gestures, the task elicited scientific argumentation as defined by NGSS in as little as 5 minutes. Because existing work shows teachers have less time for science than other subjects (Penfield & Lee, 2010), it is important to highlight possibilities for maximizing collaborative learning with limited instructional time. We posit that students' adept use of social gestures reflected earlier explicit instruction and low-stakes practice of these skills as directed by the CS manual. Students also had visual supports (e.g., anchor charts with sentence stems) of past SEL lessons available during their science discussions.

Students in comparison classrooms spent most of their time talking about logistics. Comparison group teachers tended to ask vague reflection questions and expected students to produce a physical product (e.g., a poster or worksheet) in addition to discussing science ideas. Even though students were given more time for discussion in comparison classrooms, their conversations tended to focus on the logistics of the assignment. These findings identify ways of supporting young students as they learn how to make and justify claims, use evidence, and ask questions to learn in science. They also demonstrate the necessity of social skills for groups to function. By analyzing conditions that led to rich discussions in science, we add to the growing body of literature integrating content instruction and social and emotional skills development.

Implications for Practitioners

Teacher-centered instruction dominates most typical elementary science classrooms (Reiser, 2013) with teachers positioned as the authority on what knowledge is valued or "correct." These conditions create classrooms where students are seldom challenged to think like scientists. NGSS calls for teachers to shift to a student-centered approach while also establishing boundaries that keep conversations focused. Our findings call attention to the importance of well-designed tasks, integration with social and emotional instruction, and authentic content understanding as important for elevating predominantly D-I conversations to instances of productive argumentation from evidence.

Productive, content-focused conversations occurred in classrooms where the task met three criteria: (1) a clearly articulated question with no "right" answer, (2) provision of reference materials, and (3) collaborative knowledge building as the outcome for the activity (rather than a physical product). These findings reflect prior research identifying the types of discussion activities in science that prompt students to engage in inquiry learning (Kuhn et al., 2013). When using less-defined discussion tasks, teachers should anticipate that students will have interactions about logistics. Allowing sufficient time for both negotiation and completion of an assignment allows for authentic practice using social skills while talking about science. However, with the limited instructional time available for science, that additional time might not be available. Some teachers used questions to redirect conversations, for example: "What's your evidence for that?"; "Hey Krista, what do you think?"; "Is there any one right answer?" Questions like these modeled the use of science-related argumentation practices and guided student discussions with dialogic (rather than authoritative) communication.

Limitations and Future Directions

Two limitations warrant mention and consideration in future research. First, despite a relatively large sample for qualitative research, we have limited information about the individual students in each discussion. Future work incorporating student demographics would provide a more nuanced view of their educational experiences in science. Our data were also limited to conversations from a single time point. Research suggests that neither communicative competence nor engagement in science develops linearly but in "fits and starts" (Ryu & Lombardi, 2015). The same activities in one classroom could yield different conversations on another date, even with the same student sample (Berland, 2011; Kuhn et al., 2017). Researchers could generate a more in-depth understanding of the classroom culture surrounding scientific discussions by observing multiple time points.

Concluding Comments

Students enjoy opportunities to interact with their peers. This article describes how conversations can build science knowledge and communication skills among young students. By analyzing science tasks and corresponding instruction, we identified strategies that were associated with more productive, content-relevant discussions. Sufficient foundational instruction and the right materials can bring intentionality to peer interactions so that students can use their social and emotional skills to make claims, use evidence, and ask questions to build collective science knowledge.

Note

The research reported here was supported by the Institute of Education Sciences (IES), US 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, our graduate and undergraduate students, and the collaborating district administrators, teachers, and students. Ashley Hunt is a doctoral candidate in educational psychology–applied developmental science at the University of Virginia; Sara E. Rimm-Kaufman is a professor of education at the University of Virginia; Eileen G. Merritt is a research scientist in the College of Natural Resources and Environment at Virginia Tech; Nicole Bowers is a doctoral candidate in the learning, literacies, and technologies program at Arizona State University. Correspondence may be sent to Ashley Hunt at aeh6b@virginia.edu.

References

Banilower, E. R., Smith, P. S., Malzahn, K. A., Plumley, C. L., Gordon, E. M., & Hayes, M. L. (2018). *Report of the 2018 NSSME+*. Horizon Research, Inc.

Berland, L. K. (2011). Explaining variation in how classroom communities adapt the practice of scientific argumentation. *Journal of the Learning Sciences*, **20**(4), 625–664.

Cazden, C. (2001). *Classroom discourse: The language of teaching and learning*. Heinemann.

Collaborative for Academic, Social, and Emotional Learning. (2012). *Effective social and emotional learning programs: Preschool and elementary edition*.

Collaborative for Academic, Social, and Emotional Learning. (2017). *Social and emotional learning competencies*. https://casel.org/wp-content/uploads/2019/12/CASEL-Competencies.pdf

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. https://doi.org/10.1002/sce.21243

Connect Science Team. (2017). *Connect science* [Unpublished Curriculum Manual]. University of Virginia, Arizona State University, & Harkins Consulting, LLC.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, **84**(3), 287.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, **82**(1), 405–432.

Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education*, **38**(1), 39–72.

Eccles, J. S., & Roeser, R. W. (2011). Schools as developmental contexts during adolescence. *Journal of Research on Adolescence*, **21**(1), 225–241.

Erduran, S., & Jiménez-Aleixandre, M. P. (2008). *Argumentation in science education: Perspectives from classroom-based research*. Springer.

Hackling, M., Smith, P., & Murcia, K. (2010). Talking science: Developing a discourse of inquiry. *Teaching Science*, **56**(1), 17–22.

Hanuscin, D. L., & Zangori, L. (2016). Developing practical knowledge of the Next Generation Science Standards in elementary science teacher education. *Journal of Science Teacher Education*, **27**(8), 799–818. https://doi.org/10.1007/s10972-016-9489-9

Harris, C. J., Penuel, W. R., D'Angelo, C. M., DeBarger, A. H., Gallagher, L. P., Kennedy, C. A., Cheng, B. H., & 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. https://doi.org/10.1002/tea.21263

Hayes, K. N., & Trexler, C. J. (2016). Testing predictors of instructional practice in elementary science education: The significant role of accountability. *Science Education*, **100**(2), 266–289. https://doi.org/10.1002/sce.21206

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, **42**(2), 99–107. https://doi.org/10.1080/00461520701263368

Jones, S. M., Barnes, S. P., Bailey, R., & Doolittle, E. J. (2017). Promoting social and emotional competencies in elementary school. *Future of Children*, **27**(1), 49.

Kam, C., Greenberg, M. T., & Walls, C. T. (2003). Examining the role of implementation quality in school-based prevention using the PATHS curriculum. *Prevention Science*, **4**(1), 55–63.

Kids as planners: A guide to strengthening students, schools and communities through service-learning. (2011). Harkins Consulting, LLC.

Kochenderfer-Ladd, B., & Ladd, G. W. (2016). Integrating academic and social-emotional learning in classroom interactions. In K. R. Wentzel & G. B. Ramani (Eds.), *Handbook of social influences in school contexts: Social-emotional, motivation, and cognitive outcomes* (pp. 349–366). Routledge. https://doi.org/10.4324/9781315769929.ch22

Kuhn, D., Zillmer, N., Crowell, A., & Zavala, J. (2013). Developing norms of argumentation: Metacognitive, epistemological, and social dimensions of developing argumentive competence. *Cognition and Instruction*, **31**(4), 456–496.

Lee, O., Miller, E. C., & Januszyk, R. (2014). Next Generation Science Standards: All standards, all students. *Journal of Science Teacher Education*, **25**(2), 223–233.

Lemke, J. L. (1998). Multiplying meaning: Visual and verbal semiotics in scientific text. In J. R. Martin & R. Veel (Eds.), *Reading science: Critical and functional perspectives on discourses of science* (pp. 87–113). Routledge.

Lewis, K. M., Vuchinich, S., Ji, P., Dubois, D. L., Acock, A. C., Bavarian, N., Day, J., Silverthorn, N., & Flay, B. R. (2016). Effects of the Positive Action program on indicators of positive youth development among urban youth. *Applied Developmental Science*, **20**(1), 16–28.

Manz, E. (2015). Representing student argumentation as functionally emergent from scientific activity. *Review of Educational Research*, **85**(4), 553–590.

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

National Research Council. (2012a). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. National Academies Press.

National Research Council. (2012b). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press.

National Research Council. (2015). *Guide to implementing the Next Generation Science Standards*. National Academies Press.

Nielsen, J. A. (2012). Science in discussions: An analysis of the use of science content in socioscientific discussions. *Science Education*, **96**(3), 428–456.

Pasley, J. D., Trygstad, P. J., & Banilower, E. R. (2016). *What does "implementing the NGSS" mean? Operationalizing the science practices for K–12 classrooms*. Horizon Research, Inc.

Penfield, R. D., & Lee, O. (2010). Test-based accountability: Potential benefits and pitfalls of science assessment with student diversity. *Journal of Research in Science Teaching*, **47**(1), 6–24. https://doi.org/10.1002/tea.20307

Reiser, B. J. (2013). *What professional development strategies are needed for successful implementation of the Next Generation Science Standards* [Paper presentation]. Invitational Research Symposium on Science Assessment, Washington, DC, USA.

Rimm-Kaufman, S. E., & Hulleman, C. S. (2015). SEL in elementary school settings: Identifying mechanisms that matter. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds.), *Handbook of social and emotional learning: Research and practice* (pp. 151–166). Guilford.

Rogoff, B., & Toma, C. (1997). Shared thinking: Community and institutional variations. *Discourse Processes*, **23**, 471–497. https://doi.org/10.1080/01638539709545000

Ryu, S., & Lombardi, D. (2015). Coding classroom interactions for collective and individual engagement. *Educational Psychologist*, **50**(1), 70–83.

Schonert-Reichl, K. A. (2017). Social and emotional learning and teachers. *Future of Children*, **27**(1), 137–155.

Taylor, R. D., Oberle, E., Durlak, J. A., & Weissberg, R. P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. *Child Development*, **88**(4), 1156–1171.

Tippett, C. (2009). Argumentation: The language of science. *Journal of Elementary Science Education*, **21**(1), 17–25.

Trygstad, P. J., Smith, P. S., Banilower, E. R., & Nelson, M. M. (2013). *The status of elementary science education: Are we ready for the Next Generation Science Standards?* Horizon Research, Inc.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological process*. Harvard University Press.