Improving Argumentative Writing of Sixth-Grade Adolescents Through Dialogic Inquiry of Socioscientific Issues

Jie Zhang[1], Ma. Glenda Lopez Wui[2], Rosa Nam[3], Jackie Eunjung Relyea[4] & Sissy S. Wong[1]

[1] University of Houston | USA

[2] Ateneo de Manila University | Philippines

[3] Colorado State University | USA

[4] North Carolina State University | USA

Abstract: This study investigated the effect of a four-week socioscientific issues (SSI)-based intervention on sixth-grade students' argumentative writing and transferability of argument skills across topics. Students in three treatment classrooms engaged in an SSI unit on space exploration while students in three comparable classrooms continued regular space science lessons. Argumentation skills were assessed by individual decision letters about space exploration. Argument transfer was assessed by an essay to address a novel SSI. Treatment students wrote more elaborated decision letters with stronger arguments, relied less on personal ideas, and transferred argument skills to a novel SSI after the intervention. The implications of using SSI as a promising approach to integrating science and literacy learning for diverse adolescents were discussed.

Keywords: dialogic inquiry, socioscientific issues, argumentative writing, transfer

journal of
WRITING RESEARCH

Zhang, J., Lopez Wui, Ma.G., Nam, R., Eunjung Relyea, J., & & Wong, S.S. (2023). Improving argumentative writing of sixth-grade adolescents through dialogic inquiry of socioscientific issues. *Journal of Writing Research, 14*(3), 375- 419. DOI: 10.17239/jowr-2023.14.03.03

Contact: Jie Zhang, University of Houston , Department of Curriculum & Instruction, College of Education, Houston, TX 77204-5027 | USA - jzhang64@uh.edu. Orcid: http://orcid.org/0000-0003-2223-3684

Theory and research suggest a connection between classroom discussions about complex questions and the development of argumentative thinking and writing skills in adolescents (Chen et. al., 2016; Kuhn et al., 2016; Morris et al., 2018; Reznitskaya et al., 2001). Argumentative thinking and writing skills are increasingly important in academic and non-academic life, yet students struggle to write in an argumentative or persuasive style. Only about one-quarter of adolescent students perform at or above the proficient level in writing (National Assessment of Educational Progress, 2011). In science classrooms, dialogic discourse, such as argumentation, provides students with opportunities to interact with peers' thinking processes on scientific concepts and evaluate and negotiate for collaborative meaning-making activities while navigating class communication (González-Howard et al., 2017). However, limited research is available about how peer discourse in science classrooms might support the argumentative thinking and writing of early adolescents, especially emergent bilingual (EB) students.

The primary purpose of this study was to investigate the effect of a four-week socioscientific issue (SSI)-based intervention called Dialogic Inquiry for Socioscientific and Conceptual Understanding in School Science (DISCUSS) on sixth-grade students' argumentative writing and ability to transfer these skills to different socioscientific issues. Our previous work has provided evidence of its positive effects on academic vocabulary and science content knowledge (Relyea, Zhang, Wong, Samuelson, & Wui, 2022). The current study extends the previous findings by examining how language-infused activities (e.g., reading comprehension strategies, whole-class and small-group discussions, vocabulary knowledge building, and quick write) around SSI in science classrooms can help students improve their written argumentation skills, especially for emergent bilingual students. This study was framed by three theoretical lenses: (a) dialogic interaction and argumentative writing, (b) acquisition and transfer of argumentation through socioscientific issues, and (c) argumentative writing and science learning.

1. Dialogic Interaction and Argumentative Writing

The current study drew upon Vygotsky's (1962) sociocultural theory that underscores the importance of students' engagement in collaborative dialogues for knowledge co-construction and meaning-making. Under this sociocultural view, learning is a process of personal construction through dialogue and negotiation processes over the solution to a problem or a critical issue (Kuhn, 1991; Newell et al., 2011; Resnick et al., 2015). Engaging in dialogic interaction, defined as collaborative action and dialogue among all participants in goal-directed activities, can "offer a fruitful path toward developing individual argumentative writing" (Kuhn et al., 2016, p. 3), and help students challenge their existing beliefs or offer alternative answers to a question (Doise & Mugny, 1984).

Previous empirical studies grounded in the sociocultural perspective suggest that engagement in dialogic interaction on authentic and complex issues facilitates not only adolescents' meta-level understanding of argument (Kuhn et al., 2013) but also argumentative writing skills on familiar or new topics (Kuhn & Crowell, 2011; Reznitskaya et al., 2001; Shi et al., 2019). Despite the available empirical support, the connection between classroom talk and argumentative writing in existing literature remains less impressive or under theorized. Our study aims to shed light on this connection. In this study, argumentative writing is understood as a social practice, connecting dialogic argumentation and individual argumentation. Dialogic argumentation, similar to dialogic inquiry (Wells, 1999), refers to a discourse-mediated process in which students and teachers collaborate in critically observing, examining and making claims, offering evidence, providing reasoning, making counterclaims, and offering alternative explanations (Crowell & Kuhn, 2014). We share the stance that dialogic argumentation is a productive bridge to individual argumentative writing (Kuhn 2015). Quality of argumentative writing in the current study is conceptualized as "identifying and weighing positive and negative attributes of contrasting positions on the issue, drawing on relevant evidence to inform the judgments involved" (Kuhn & Crowell, 2011, p. 546).

Dialogic interaction is important in content area teaching given the increasing linguistic demands of the Common Core Curriculum Standards. For example, the Next Generation Science Standards (NGSS) call for a shift in traditional teacher-centered science teaching by highlighting the importance of guiding students to understand and explain phenomena with a set of science and engineering practices (National Research Council, 2015). This change involves the emphasis on students' engagement in higher-level reasoning and discussions and sophisticated and language-intensive science practices. Recent research has shown the positive effects of dialogic interactions on students' written scientific arguments (McNeill & Krajcik, 2008; Murphy et al., 2018). High school students who received an intervention called Quality Talk Science (QTs) in which teachers and students learned to use productive discussions in six QTs-enhanced science lessons were more able to provide evidence-based written scientific arguments than comparison students (Murphy et al., 2018). Indicators of productive discussions include high-level thinking, speculation, uptake, connection, and personal experience (Soter et al., 2008), and different types of talk (i.e., nontransactive talk, cumulative talk, and exploratory talk) (Mercer, 1996; Zhang et al., 2016). Nontransactive talk refers to short exchanges such as simple explanations, cumulative talk involves mainly confirmations and explanations, and exploratory talk entails constructive dialogues and critical engagement, such as counterarguments, rebuttals, and high-level questions.

Despite the importance of dialogic interaction in science learning, developing argumentation skills through interaction does not occur automatically and requires explicit instruction and sustained and routine discourse practices, particularly for bilingual students to achieve the simultaneous development of science content knowledge, English language proficiency, and argumentation skills (González-Howard & McNeil, 2019; Osborne et al., 2004; Zohar & Nemet, 2002). Simple participation in dialogic interactions about controversial issues may not translate to immediate student learning gains, especially for bilingual students, because it takes time for teachers and students to develop discourse practices that promote higher-order thinking and argumentation skills in science classrooms (Murphy et al., 2018), and students need additional scaffolds to fully participate in the classroom discourse. However, it is beyond the scope of this paper to discuss teacher scaffolding and teacher changes in beliefs and attitudes.

2. Socioscientific Issues (SSI) and Argumentation

The current SSI-based DISCUSS approach to learning scientific argumentation emphasizes the interaction between science and society, including moral, ethical, and political influences on decision-making in scientific contexts (Cavagnetto, 2010). Socioscientific issues are controversial social issues with conceptual or procedural links to science and contentious issues without clear-cut resolutions (Sadler, 2004) and are "open-ended, ill-structured, debatable problems subject to multiple perspectives and solutions" (Sadler & Zeidler, 2005, p. 113). SSI-based instruction implemented in science classrooms has shown positive effects in promoting scientific reasoning, argumentation, decision-making, and science content learning for students from various cultures and grade levels (See review, in Sadler, 2011). SSI-based instruction has also shown promise in improving adolescent students' quality of argumentation in oral and written domains (Atabey & Topcu, 2017; Chowning et al., 2012; Dolan et al., 2009; Gutierez, 2015; Venville & Dawson, 2010; Zohar & Nemet, 2002), but how culturally and linguistically diverse students engage in SSI and its potential benefits on language and thinking development is less understood.

The positive impact of SSI-based intervention on written argumentation has mostly been documented among high school students (e.g., Dawson & Carson, 2020; Venville & Dawson, 2010) and reported only in a few studies involving upper elementary and middle school students (Atabey & Topcu 2017; Belland et al., 2015; Morris et al., 2018; Zhang et al., 2016). For example, Atabey and Topcu (2017) found pre- and post-test improvements in 7th graders' argumentative writing in terms of the quality of the claim, evidence, and reasoning components after the implementation of an SSI-based curriculum on global warming in science classrooms.

Going beyond the previous research, the primary aim of the current study was to examine the impact of an SSI-based curriculum on argumentative writing in sixth-grade science classrooms in an urban middle school with a large number of linguistically diverse students. The study focuses on early adolescents. Sixth grade, in particular, is important because it is a critical period when science instruction and content become more in-depth, students start to circumscribe their career interests (Turner & Lapan, 2005), and student attitudes, especially for girls toward science, generally decline (George, 2000).

3. Acquisition and Transfer of Argumentation Skills

The second aim of the study was to investigate whether the acquired argumentation skills through the curriculum-based SSI would transfer to a novel SSI. Transfer is the application of intellectual skills, acquired in a specific context, to other situations that may be similar or different (Foong & Daniel, 2013; Khishfe, 2013). In the current study, the transfer of argumentative skills refers to the student's ability to apply the argumentation skills learned from participating in an SSI-based instruction to reason about a new SSI that is not discussed during the intervention period. Despite the positive findings of SSI, only a few studies have explored the possibility of transfer of argumentation skills from SSI-based instruction to other real-world issues ranging from the moral dilemmas of everyday life (Herman et al., 2021; Iordanou, 2010; Zhang et al., 2016; Zohar & Nemet, 2002) to other controversial social issues involving the application of scientific knowledge (Foong & Daniel, 2013; Khishfe, 2014).

In an eight-week SSI-based intervention on water usage and safety, Khishfe's (2014) study found that seventh graders received explicit argumentation instruction with detailed lessons on argumentation and justification processes. The results showed the transfer effect of argumentation skills by comparing treatment and control group students' argumentative writing in pre- and post-tests. The pre-test writing prompt was related to the primary topic of the unit (i.e., *Would you vote for adding Fluoride to drinking water in your city?*), while the post-test prompt was a new SSI topic around genetically modified food that was not addressed during the intervention (i.e., *Do you think the golden rice should be produced and marketed?*). Students who participated in the treatment condition demonstrated significantly stronger argumentation skills in responding to the transfer writing prompt.

In another study that tested the effects of SSI-based intervention for ninth-grade students called the Genetic Revolution unit classroom discourse, Zohar and Nemet (2002) found that a 12-hour intervention unit designed to exercise the principles of good argumentation (having multiple, true, and factual justifications and providing rebuttals against counter-arguments) through written tasks and group discussion improved the quality of the students' argument writing, particularly in the number of justifications and argument structure. The authors argue that the success of the intervention was attributed to direct instruction in argumentation skills and multiple opportunities to engage in discussion on biological concepts, as well as that the unit built upon and stimulated students' preexisting skills. In addition, students were able to transfer the reasoning abilities taught in the context of human genetics to the context of dilemmas taken from everyday life.

In a large-scale intervention study involving predominantly African-American and Hispanic fifth-grade students, the findings showed positive transfer effects of an SSI-based decision-making curriculum involving peer-led small group discussions, called Collaborative Reasoning (CR), and collaborative work on argumentation skills across different topics (Morris et al., 2018; Zhang et al., 2016). The discussion approach, CR, is a dialogic approach to the small group discussion that calls for critical and reflective thinking (Clark et al., 2003). For the CR discussion, teachers detailed the ground norms and thereafter posed the Big Question and invited students to participate in the discussion. Students were expected to manage their discussions and state their positions, come up with supporting evidence, challenge their peers, and consider opposing opinions. Treatment students who participated in a multidisciplinary unit, featuring small group CR discussions and collaborative work on a controversial community issue and integrating environmental science and public policy, were more able to produce elaborated reasons and consider both sides of the issue than the students in the direct instruction condition (Morris et al., 2018). More interestingly, such intervention effects were transferred to a novel-writing task. In reflective essays on a topic unrelated to the issue covered in the intervention curriculum, the treatment group students were better able to use a comprehensive set of reasons in support of their arguments, recognize different sides of the issue, and weigh the importance of reasons compared to the students in the direct instruction condition (Zhang et al., 2016).

The existing literature (see a review, by Resnick et al., 2018; Engle et al., 2011; 2012) has proposed several possible explanations for the mechanism underlying argument transfer: cognitive conflict (Adey & Shayer, 2015), sociocognitive (Chan et al., 1997; Chi & Wylie, 2014), thinking dispositions (Zohar & Nemet, 2002), proactive executive control (Nussbaum & Asterhan, 2016), metacognitive (Koedinger & Wiese, 2015; Reznitskaya et al., 2008), motivational-social (Dweck, 2006), and framing contexts (Engle et al., 2011; 2012). Some scholars propose that the cognitive mechanism of argument skill transfer is enhanced by combining cognitive conflict with social/peer interaction (Chan et al., 1997; Chi & Wylie, 2014). Students facing cognitive conflict benefit from peer collaboration in knowledge processing exhibited as constructive and extended joint dialogues. Transfer is more likely to happen when the learning context is framed expansively as opportunities for students to engage in larger conversations that are not bounded by times, locations, people, and activities (Engle et al., 2011; 2012).

Other scholars posit that the structure of dialogic talk and change in what is valued in classroom dialogues shape what students think is expected of them and will change students' thinking dispositions and patterns of responses (Zohar &

Nemet, 2002). Nussbaum and Asterhan (2016) further elaborate on this view and propose that participation in argumentative discourse strengthens proactive executive control strategies because participation in argumentation involves active consideration of other's counterargument, inhibiting interference from own argument, and attention shift between different sides of the argument. At the *metacognition* level, when students develop a repertoire of reasoning skills through dialogic interactions (Koedinger & Wiese, 2015), they gradually internalize an argument schema, an abstract knowledge structure about the components of a complete and sound argument, through participation in argumentative discourse (Reznitskaya et al., 2008). Explicit instruction of argument structure may contribute to students' acquisition of arguments on a metacognitive level — being conscious of the generalizations, principles, and standards of one's reasoning processes (Zohar & Nemet, 2002). We adopt this line of theoretical underpinnings — fostering student thinking dispositions, patterns of responses, and metacognition by creating space for open participation and dialogic interactions —in our curriculum design and instruction.

4. Argumentative Writing and Science Learning in Bilingual Science Classrooms

Moving away from using writing as a tool for students to demonstrate their knowledge, recent research on writing to learn science has used writing as a tool to construct knowledge while conducting science investigations (Ardasheva et al., 2015). Most writing tasks in science classrooms include conventional tasks (i.e., laboratory reports or science explanations), reflective writing (i.e., reflective journals), or a mix of both formats. To date, limited research is available on emergent bilingual students' argumentative writing in science. Huerta and Garza (2019) reviewed the use of writing in science and literacy-integrated interventions with both emergent bilingual students and native English speakers from 1996 to 2016. The majority of writing interventions in science reported positive learner (including bilingual students) outcomes in terms of conceptual understanding and academic language (Garza et al., 2018; Lee et al., 2009). However, fewer studies have focused on emergent bilingual students' argumentative writing compared with native English-speaking students (e.g., Cervetti et al., 2012; Hand et al. 2016). Most science writing research with emergent bilingual students has focused on abilities to use science language forms (e.g., content vocabulary) to reflect understanding (de Oliveira & Lan 2014; Kim & Kim, 2021; Lee et al., 2009; 2011), and learners' ability to transfer learning to measures of science vocabulary and science reading comprehension (Cervetti et al., 2012; Lara-Alecio et al., 2012).

5. The Present Study

Building upon previous work on dialogic argumentation and writing in science, the current study aimed to investigate sixth-grade students' experiences with SSI and argumentative writing in two general education classrooms and one bilingual science classroom, in comparison with three comparable business-as-usual classrooms that did not receive the SSI unit or argumentative writing instruction. The DISCUSS Curriculum features a language-infused, SSI-based unit on Space Exploration. Pre-post gains in written argumentation skills were assessed by individual decision letters about space exploration in the treatment group. As a post-intervention measure, both treatment and comparison groups wrote an essay to address a novel SSI. More detailed descriptions of the DISCUSS intervention are available in the Methods section.

Three research questions (RQs) guiding the current study and corresponding hypotheses are as follows:

1. To what extent does treatment students' argumentative writing improve from before to after the SSI-based DISCUSS curriculum?

We hypothesized that students who receive the DISCUSS treatment will improve their argumentative writing skills, particularly in the claim, evidence, and reasoning framework before and after the intervention.

2. To what extent do treatment students transfer the argumentation skills acquired during the DISCUSS intervention to solve a novel socioscientific issue, and demonstrate improved outcomes when compared with a comparison group?

We hypothesized that the students who participate in the DISCUSS curriculum will apply the acquired argumentation skills to reason about another novel SSI, and write better argumentative essays than the students in the comparison condition. Dialogic interaction practices can help students to develop a repertoire of metacognitive and reasoning skills so that students gradually internalize an argument schema that can be applied and utilized in a new task.

3. How do students' talk patterns relate to their argumentative writing?

Based on prior literature, we hypothesized that the indicators of productive discussion (e.g., questioning, evidence, exploratory talk.) are positively correlated with the reasoning quality of student writing. Students who participate more in co-constructed dialogues would write better argumentative letters or essays.

6. Method

6.1 Participants

Participants were 137 sixth graders (females 49%) from six classes at an intermediate school located in an urban district in southeastern Texas. The district's student population is composed of 54.2% Hispanic, 28.9% African-American, 11.7% Asian, 3.8% White, and 2.4% others. Additionally, 84.4% of the students are identified as economically disadvantaged and 43.7% as emergent bilinguals (EBs) or commonly known as English learners (ELs). Over 80 languages are represented in the district with 75% of ELs being Spanish speakers; the other prevalent languages spoken are Vietnamese, Urdu, Yoruba, Burmese, and Arabic. In the current school, about 80% of the students were eligible for free or reduced-price lunches (The Texas Tribune, 2017-18).

Among the 137 students, 58 (42%) were Hispanic, 36 (26%) were African American, and 12 (9%) were Asian. Of the sample, 63 students (46%) were native English speakers, and 74 (54%) students were non-native English speakers. Of these, 58 students (78%) spoke Spanish and other languages spoken were Thai, Malay, Vietnamese, Burmese, Kareni, Filipino, Arabic, and Portuguese. According to a parent survey, 33 out of 74 (43.7%) ELs were born in the United States.

6.2 Research Design and Procedure

The study used a mixed-methods, pretest-posttest study combined with a post hoc comparison with a no-treatment group. Three classrooms (n = 73; 36 girls, 37 boys), including two general education classrooms serving diverse students and one Spanish-English bilingual classroom serving all Spanish-speaking bilingual students (2 newcomers), were randomly assigned to the treatment group and received the SSI-based DISCUSS curriculum. The other three comparable classrooms ($n = 64$; 31 girls, 33 boys), two general education classes, and one bilingual class served as the no-treatment comparison group and received the Space Science unit following the district science sequence and pacing guide. Students in the treatment condition participated in a series of small-group discussions and one target group of students (4-5) in each classroom was videotaped throughout the intervention period. The target group was a representative cross-section of each treatment classroom in terms of gender, ethnicity, talkativeness, and academic achievement.

Teacher Professional Development.

Three treatment group teachers (Ms. J, Ms. D, and Ms. L) participated in a full-day workshop before the intervention implementation. Ms. J and Ms. D taught the general education classes and Ms. L taught the bilingual class. Teachers reviewed the drafts of the DISCUSS curriculum and provided the research team with feedback before the workshop. During the PD workshop, teachers were introduced to the theoretical and research background of the study, and reviewed the finalized curriculum materials and lesson layout. The research team modeled the key science strategies: inquiry, Claim-Evidence-Reasoning (CER), engineering design (stomp rocket system kit), and literacy strategies (academic vocabulary, reading comprehension, graphic organizer, small group CR discussions, and scaffolded writing).

DISCUSS Intervention

Our research team partnered with a group of sixth-grade science teachers to develop and implement a four-week unit on Space Exploration. The unit addressed a contestable question: *Should the U.S. government increase or decrease funding for space exploration*? The curriculum featured a 7E (**Elicit,** Engage, **Establish,** Explore, Explain, Elaborate and Evaluate) instructional model (August et al., 2014), based on the Biological Science Curriculum Study's (BSCS) 5E model, to address student background knowledge (Elicit) and vocabulary (Establish), the use of classroom discussions and the CER framework (McNeil & Krajcik, 2012) to facilitate reasoning and argumentation, and conceptual understanding in space science lessons. In week one lesson one, students were exposed to the SSI and read the newsletter that introduced real-world issues related to Space Exploration, and explored their pre-conceived notions about the central issue: *Should the government increase or decrease the funding for space exploration*? In the following four-week period, students read an argumentative text each week and engaged in small-group group discussions about the impact of space exploration on technological innovation, earth and space environment, economy, and public policy. Figure 1 illustrates the curriculum layout.

Figure 1. Space Exploration Unit Curriculum Texts.

Four argument texts associated with space exploration issues—technology, environment, economy, and public policy—were written by the research team to provide students with balanced arguments about the pros and cons of space exploration. Figure 2 illustrates the structure of a sample argument text. Throughout the argument texts, clear headings were used to highlight each side of the argument; key science content-specific vocabulary (e.g., emissions, degradation) and general academic vocabulary (e.g., insulate, resilience) were defined; in addition, interesting and relevant background information and visual representations (e.g., graphs and pictures) were provided to facilitate text comprehension.

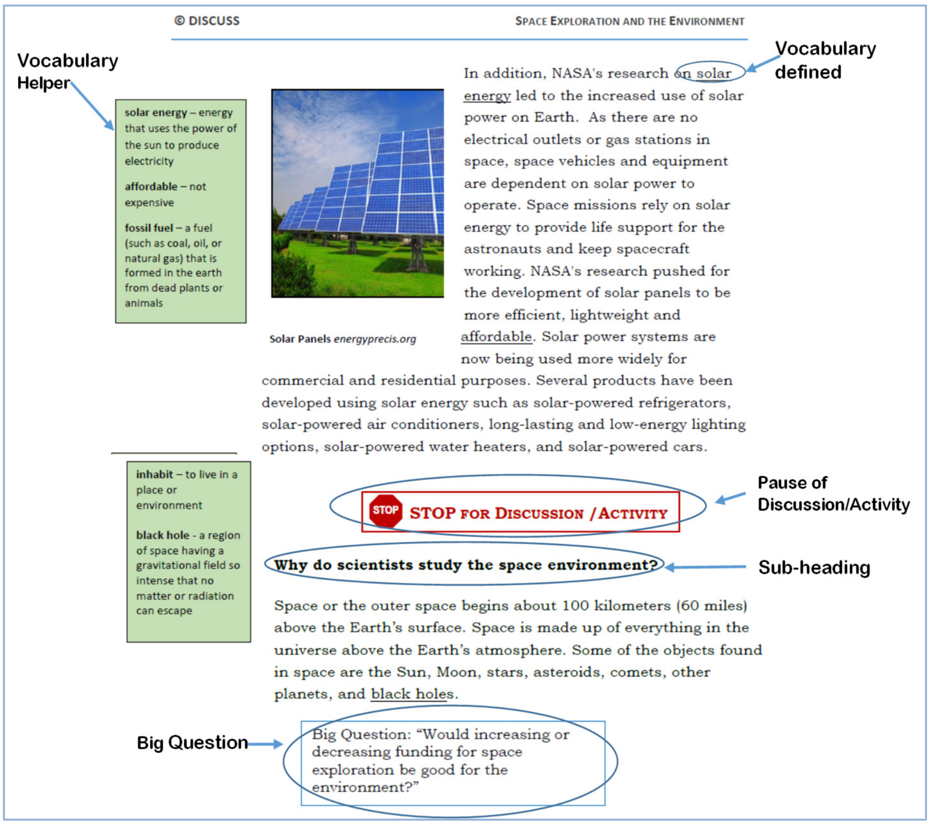

Figure 2. Structure of a Sample Argument Text.

Each week, students in the treatment condition read one argument text and engaged in small-group discussions and essay writing considering different perspectives. After reading the text, the class was divided into groups of 5-7 to engage in Collaborative Reasoning discussions about the Big Question listed below.

- Week 1 – Technology: "Would increasing or decreasing funding for space exploration be good for technological innovation?"
- Week 2 – Economy: "Would increasing or decreasing funding for space exploration be good for the economy of the United States?"
- Week 3 – Environment: "Would increasing or decreasing funding for space exploration be good for the environment?"
- Week 4 – Public Policy: "Would increasing or decreasing funding for space exploration be the best public policy for the American people?"

To start the CR discussion, teachers first set up the ground norms, then posed the Big Question and invited open participation. Students were encouraged to manage their discussion and express their positions/claims, provide supporting evidence, challenge one another, and consider alternative perspectives. Teachers used a series of scaffolding moves including prompting for position and reasons, modeling and thinking out loud, asking for clarification, challenging, reminding, encouraging, fostering independence, summing up and re-focusing, and debriefing (Clark et al., 2003). The CR discussions typically last about 10 minutes.

After the CR discussion, each group was asked to summarize their main arguments and then create an argument diagram. The students wrote their claims, evidence, and reasoning on a large post-it chart. Students were encouraged to write arguments for both sides of the issue, in favor of increasing and in favor of decreasing columns of the chart. Teachers were provided an argument outline illustrating the main arguments on both sides. After completing the argument diagram, students engaged in quick write to address the Big Question using the CER framework (McNeill & Krajcik, 2011). On the last two days of the unit, students reviewed the key arguments on all four domains and evaluated the strengths and weaknesses of the arguments within and across domains. Students then engaged in another CR discussion about the central question. After the discussion, students wrote an individual decision letter addressed to the Space Exploration Agency expressing their recommendation on whether to increase or decrease funding for space exploration. During the intervention, research assistants conducted classroom observations, took field notes, and helped with video recording the lessons.

No-treatment Comparison Group

Three other comparable classrooms—two general education classrooms and one bilingual classroom—served as the no-treatment comparison group. The comparison classes followed the district's sixth-grade science scope and sequence pacing guides and continued business-as-usual science lessons for the *Earth and Science* unit. Overall, the lessons followed the 5E science instructional model as required by the district, and included hands-on components in some form. For example, one lesson had students ball up paper to model the different sizes of the planets in our solar system. What was noticeable in the comparison classes when compared with the experimental classes was the lack of integration of language and literacy instruction. In general, class discussions either in a whole or small group regarding the scientific concepts, investigations, and student ideas rarely occurred (less than 5 minutes in a lesson, if any). A few types of writing activities existed as a form of answering literal and inferential comprehension questions, there were limited opportunities for students to engage in writing to reflect their science learning and argumentation.

Measures

Before the intervention, all classes were administered language and science background measures: Test of Silent Word Reading Fluency, science content knowledge, and science academic vocabulary. After the intervention, all students were assessed on academic vocabulary, argumentative writing, science content knowledge, and science engagement. Due to the scope of the study, only argumentative writing outcomes will be reported in the current study.

Pre-Intervention Measures

Word reading fluency. The standardized, normed-reference Test of Silent Word Reading Fluency (TOSWRF) (Mather et al., 2004) was used. Students were provided with rows of unrelated words of increasing difficulty with no spaces separating them (e.g., *dimhowfigblue*) and given three minutes to draw lines between as many words as they can (e.g., *dim|how|fig|blue).* The test-retest reliability of the instrument is .92 (Mather et al., 2004).

Science content knowledge

Students' science content knowledge was measured using a researcher-developed test that specifically addressed the four main theme areas of space science: solar system, planets, gravity, and rocketry. The test included multiple-choice and matching questions. The internal consistency reliability of 30 items was .74.

Academic vocabulary

A researcher-designed measure was administered to assess students' academic vocabulary knowledge related to space science. The test contained 40 multiple-choice questions to choose a synonym for a given word. A total of 40 academic vocabulary words (20 general academic and 20 science-specific vocabulary words) were randomly selected from the target vocabulary words taught throughout the Space Exploration unit and derived from the argument text booklet. The test demonstrated acceptable internal consistency reliabilities for the pretest (Cronbach's alpha = .73).

Post-intervention Measures

The current study focused on two data sources of argumentative writing: pre-post decision letters on the space exploration budget issue for the treatment group, and post-unit argumentative writing on a new prompt for both the treatment and comparison groups. The writing tasks are described below.

Space exploration decision letter

This is a curriculum-embedded assessment. Students in the treatment condition wrote an individual letter to express their decisions on the central question during the first and last lessons. No-treatment comparison students were not asked to write the letter because they were not exposed to the space exploration issues in the Introduction packet of DISCUSS curriculum. Students were given 20 minutes to write on a letter template.

Antarctic research essay

This task was to assess student knowledge of components of a complete and sound argument by addressing a novel problem similar to the space exploration issue. Students in both the treatment and comparison classes first read a 500-word argumentative essay about the pros and cons of research in Antarctica and then completed a transfer essay on research in Antarctica. To facilitate comprehension, students listened to the audio recording of the text while reading along. Students then wrote an essay to respond to the central question: *Should we fund research in Antarctica?* The space exploration issue and Antarctic research shared similarities in the deep structure of the nature of the funding allocation problem and multiple points of view. Students were given 25 minutes to write.

Coding

Decision letter

All the pre- and post-intervention decision letters were coded using NVivo 12 by the first two authors. The coding involved three stages. First, the letters were chunked into idea units, which refer to "expresses one action or event or state, and generally corresponds to a single verb clause" (Mayer, 1985, p. 71). Next, each idea unit was coded using the coding scheme in Table 1. The coding scheme was adapted from Kuhn and Crowell (2011) and Morris et al. (2018). Three broad categories were generated from the code: pre-conceived notions, domain reasons, argument quality, and type. Preconceived notions are based on student's prior knowledge, beliefs, and experience. Domain reasons included technology, economy, environment (space and earth), and public policy. Total scores (number

Table 1. Definitions and Coding Examples of Space Exploration Decision Letters

Code	Definition	Example
Claim	Position on the big question	The Government should increase/decrease space exploration funding.
Domain Reason		
Technology	Innovations	Space exploration has resulted in a number of technological innovations. / The Apollo space program research gave birth to micro-computers /which led to the invention of laptops, tablets, cell phones, and the GPS (Global Positioning System). /
Economy	Job creation	I also think they should increase funding for space exploration /because you need more workers./ More workers equal more money./
Environment	Space junk	But when people get sent to space the space shuttle loses parts that's called trash /when it becomes a part of space and then the trash comes back to earth /that can kill people because of how fast it comes and how big it is. /
Public Policy	Scientist view	I also think that it can help us find cures for people who have a bad disease./
	Military official view	The military is involved in providing assistance to people /when there are disasters. /Space satellites enable members of the military to communicate with each other during disasters./
Preconceived Notion	Personal beliefs and assumptions	NASA has too much money and instead could come to us giving us a better chance of having a better education. / I think it is very dangerous to go to space cause anything can happen./
Evidence	Textual evidence	It was estimated that about 1,200 died as a result of the hurricane./
Elaborated Reason	Information increases the relevance or further justifies an explicitly stated reason.	And the reason why I think they should increase is /because the military, the military can help us with a lot of things / like a rescue mission, disasters, descending our country, /like this example when the hurricane hit the gulf coast area like Louisiana, Florida, and the Bahamas…

Multi-link Reasoning	Organize incoming information and bridging inferences into coherent causal chains.	Then there's the mining effect, / when mining is involved/ it causes harm to earth /by spreading chemicals during explosions and causing air pollution./ Steel is a basic material for space equipment. /In order to get steel we need to extract it from the earth such as mining./
Argument Type		
No argument	No clear reason is given	I think the US government should increase the funding for Space Exploration/ because it's worth it/ because they are going to explore something that is dangerous/
One side only	Include only positives of a preferred option	I think the US government should increase the funding for space explorations/ because, they have enormously contributed to technological innovations that have improved our way of living /and also because, airplanes can become safer /because of the fire-resistant materials developed by NASA. /In addition NASA's educational programs for people. /
Dual perspectives	Include negatives of other options	The US government should increase funding for space exploration/ because…. But if they decrease it, even the children's parents can pay for he/her to go to the space exploration center. / That is why I think they should increase it. /
Integrative perspectives	Include negatives of a preferred option or positives of other options	The US government should decrease /because some mission could fail! /Pollution in the river that was caused by mercury... /It is true that it gave education for kids and more lessons /but how do they get that information from? /Sending aircraft and most are failing for kids who want to learn!

of idea units) coded for pre-conceived notions and domain reasons were calculated, respectively. Each letter was coded for the presence of argument quality indicators including elaborated reasons, evidence, evaluation, and multi-link reasoning. Argument type (no argument, one-sided argument, dual perspectives, and integrated perspectives) was holistically coded for each letter. First, the first two authors practiced coding a few letters and compared their codes for each letter. In cases where there were differences in coding, they went back to examine the coding scheme and arrived at a consensus on what codes to adopt. Then the first two authors coded the letters independently. The percent agreement of independent coding between the two coders ranged from 82% to 99% across individual codes.

Table 2. Definitions and Coding Examples of Antarctic Essays

Code	**Definition**	**Example**
Claim	Position on the big question	We should/should not fund research in Antarctica.
Domain Reason		
In favor of funding	Understand and stop global warming	Funding for Antarctica will improve our knowledge of global warming./
Against funding	Invest in other scientific problems	The funds for Antarctica research should be used for better investigations/ like studying sea turtles nesting sites off the gulf coast. /
Evidence	Textual evidence	Data from space satellites show/ that Antarctica has been losing more than one hundred cubic kilometers of ice each year since 2002./
Elaborated Reason	Information increases the relevance or further justifies an explicitly stated reason.	I think that we should fund research in Antarctica. / My first reason why we should fund research is to stop global warming./ For example, if the ice caps keep melting, global sea levels rise for cities like New York, Miami, and Galveston. / An additional reason why is for the ice caps to stop melting…

Multi-link Reasoning	Organize incoming information and bridge inferences into coherent causal chains.	We should fund research for Antarctica cause,/ Antarctica ice sheets are melting rapidly at a hundred cubic kilometers per year./ The massive loss of ice sheet rises the sea level./ When sea level rises/ it causes huge flooding./ If we don't try to find a solution for global warming/ and continue with our daily lives /the melting rate in Antarctica with increase./ The faster the ice sheets melt,/ the faster sea level will rise,/ causing flooding problems all over the world./ The flooding would cause chaos and war for land …
Argument Type		
No argument	No clear reason/contradictory reason is given	I think that we should fund research for Antarctica /because they're already doing what they're trying to do in Antarctica at the north pole./They should focus on endangered species instead of funding research in Antarctica./
One side only	Include only positives of a preferred option	Funding for Antarctica will improve our knowledge of global warming/, which is the warming of the atmosphere caused by excessive carbon dioxide emission/common sources of carbon dioxide emission includes cars, factories, and ranching. / The Antarctic ocean is important to stabilize the global climate because it absorbs carbon dioxide./ It is important to study and monitor the condition of the Antarctic ice sheet./
Dual perspectives	Include negatives of other options	I think we should fund research in Antarctica / because scientists don't know what's happening in Antarctica. … If scientists don't study Antarctica humans who lived there can die / because they can lose carbon dioxide one day from the

		Antarctic ocean and they wouldn't know what to do. /
Integrative perspectives	Include negatives of a preferred option or positives of other options	I think we should fund research in Antarctica / because if we know more about Antarctica and if it does for some reason cause a flood in coastal cities like New York, and Galveston … We may need to decrease a little bit / if we start having more and more endangered species / because basically without animals some things won't work the same./

Antarctic essay

A coding scheme (Table 2) similar to the decision letters was developed and the same coding process was applied to the Antarctic essays. Two broad categories were generated: domain reasons and argument quality/type. Domain reasons included reasons in favor of or against increasing funding for Antarctic research. Specific reasons are listed in Table 2. Total scores (number of idea units coded) for domain reasons were calculated. Each letter was coded for the presence of each argument quality indicator including the elaborated reasons, evidence, evaluation, and multi-link reasoning. Argument type (no argument, one-sided argument, dual perspectives, and integrated perspectives) was holistically coded for each essay. The first two authors coded the essays independently. The percent agreement of independent coding between the coders ranged from 78% to 99% across individual codes.

Target group discussion

One target group of 4-6 students from each treatment class was videotaped throughout the intervention period. A total of nine videotaped lessons were transcribed for the three target groups in the treatment classes. Three discourse-rich lessons in which students engaged in small group discussions were selected for each target group. Specifically, the first lesson and fifth lessons in week one and another SSI lesson in week three or four were selected. Each selected target group lesson was transcribed and coded.

Following McNeill and Pimentel's (2010) and Zhang, Niu, Munawar, and Anderson's (2016) coding scheme, target group student talk was coded on argument structure and dialogic interaction (or dialogic aspect of argumentation). Each student speaking turn was coded as a type of argument structure and chunks of speaking turns were classified into each type of dialogic interaction. The argument structure included the claim, question, evidence, reasoning, and rebuttal. Dialogic interaction included independent, connected, dismissal, and acknowledgment. Dialogic interaction was further coded into three-level talk: non-transactional talk (independent), cumulative talk (acknowledgment and confirmation), and exploratory talk which involves counterarguments and challenges. The detailed coding scheme is provided in Table 3. The second author coded all student talk and the third and fourth authors were trained to use the same coding scheme to independently code half of the student talk. The percent agreement of independent coding between the second and third authors ranged from 83% to 100%, and the percent agreement between the second and fourth authors ranged from 85% to 100% across individual codes.

Table 3. Coding Scheme for Student Talk: Argument Structure and Discourse Feature (Adapted from McNeill & Pimentel, 2010; Zhang, Niu, Munawar, & Anderson, 2016)

Argument structure	**Definition**	**Example**
Claim	Position about the big question	I say that um increasing funding for space exploration would be good for technological innovation because…
Evidence	Data either in support of or against space exploration. The evidence includes scientific data, personal experiences or textual information.	I think, I think we should increase because … it would be very difficult to conduct rescue operations without satellites. That saying, with the furthering we have made with space exploration, we now have satellites [textual evidence].
Reasoning	Justification for why the evidence supports the claim.	I think that they should increase it because you know when um in outer space, they're going to explore more things and the more stuff they explore the more people will get interested, and the more people are interested the more people work there and the more people that work there, the more the people that will help to build the things ...
Question	Question about the discussion	Why do you think that we should increase the funding like why do you think that?
Rebuttal	Providing counter-argument	But I think they should also increase it so we could have actual true facts about like the sun and all the planets. Because we're just estimating that information.

Discourse feature	Definition	Example
Nontransactive Talk	Short exchanges consisting of assertions and counter-assertions; not connected to previous utterance	Teacher: What does it mean to explore space? Turn to your neighbor and talk to them about what it means to explore space. Student 1: Exploring space. Student 2: Explore. Student 1: Exploring space. Student 2: It's being curious!
Cumulative Talk	Students take over, integrate and apply the perspectives of their peers, but without real disagreement or constructive conflict; characterized by confirmations, expansions, and elaborations.	Student 1: How technology affects the way we plan for articles? Student 2: So basically global warming it, like there's technology for global warming it says when it's going to happen and how it's going to happen? Student 1: How environmental technology has affected how ... it actually has helped us to tell us what's going to happen after like it helps us prepare us like it helped us to know when Harvey was going to hit! It helped us know when Harvey was going to hit, it helped us know how hard the wind was going to hit
Exploratory Talk	Students challenge or counter-challenge while justify challenges and offer alternative perspectives; characterized by constructive and critical engagements.	Student 1: I agree with both because some people are losing money as well. Like how are they going to pay their family? Student 2: I think the U.S. government should increase it, though. Student 1: But I think they should also increase it so we could have actual true facts about like the sun and all the planets. Because we're just estimating that information. Student 2: I think they should increase it. Student 1: I think they should do both.

Data Analysis

To answer the first research question, treatment students' pre- and post-intervention decision letters were analyzed for the quality and quantity of arguments. The quantity of writing was measured by the essay lengths (total words and the total number of idea units). The quality of arguments was measured by the number of idea units coded as domain reasons and preconceived notions, and the number of letters that exhibit each argument type (no argument, one-sided only, dual perspective and integrative perspective) and quality indicators (e.g., elaborated reasons, and multi-link reasoning). Non-parametric statistics were used to compare pre-post letter lengths (Related-Samples Wilcoxon Signed Rank Test).

Chi-square statistics were used to compare pre-post differences in argument quality measures.

To answer the second research question, we first compared the treatment group's post-letters and post-Antarctic essays, and then compared treatment and comparison group differences in post-Antarctic essays. The same essay length and argument quality measures as the decision letters applied to the Antarctic essays. Non-parametric statistics were used to compare treatment vs. comparison differences in Antarctic essay lengths (Independent-Samples Mann-Whitney U Test). Chi-square tests were used to compare the treatment vs. comparison differences in argument quality measures.

To answer the third question, Pearson correlations of the target group student discourse data (productive talk indicators) and their argumentative writing performance were calculated. Qualitative discussion excerpts and student writing samples were presented to illustrate the connection between the type of talk and student writing outcomes for the target students in the treatment classes.

7. Results

Table 4 presents descriptive statistics and independent samples t-test results of pretest background measures by treatment and comparison groups. The results showed no significant difference between treatment and comparison groups on three pretests (Test of Silent Word Reading Fluency, science academic vocabulary, and science content knowledge), suggesting the equivalence of treatment and comparison classes on science, reading, and vocabulary knowledge, *ps* > .05.

Table 4. Descriptive Statistics of Pretest Measures between the Treatment and Comparison Groups

	Treatment Group (n=73)		Comparison Group (n=64)				
	M	*SD*	*M*	*SD*	*T*	*p value*	*Cohen's d*
TOSWRF Raw	84.11	29.53	86.66	34.81	-.42	.67	-.08
TOSWRF Standard Scores	90.58	14.68	92.37	18.03	-.58	.55	-.11
Academic Vocabulary	17.03	6.38	17.91	5.67	-.81	.41	-.14
Science Content Knowledge	11.36	4.14	9.83	5.75	1.74	.08	.30

RQ1: Treatment Group's Pre-post Gains in Decision Letters

To answer RQ1, the pre- and post-intervention decision letters on the space exploration budget written by the treatment group were analyzed for writing length, domain reasons, and argument type and quality.

Writing length

Table 5 presents the descriptive statistics of length measures of pre-post decision letters: total words and total idea units. Related-Samples Wilcoxon Signed Rank Tests suggested that students wrote significantly more words, $Z = 1460$, $p < .001$, and more idea units, $Z = 1152$, $p < .001$ in the post-letters than in the pre-letters.

Domain Reasons

Table 6 presents the frequencies and percentages of students that had none, one, or more than one reason in the four domains and preconceived notions. The results showed that compared to the pre-letters, more students gave more reasons and far fewer students gave no reasons in the post-letters. Specifically, overall the number of letters containing no domain reasons dropped three times from pre- (64.6%) to post-letters (26.2%), but the letters with at least one domain reason doubled from pre- (35.4%) to post-letters (73.8%). In terms of the sources of reasoning, students drew the most reasons from technology perspectives, followed by public policy and environmental points of view. Economy perspectives were mentioned the least. Compared to the pre-letters, students drew significantly more reasons from technology aspects, χ^2 (2) = 9.97, $p = .007$; and had significantly fewer preconceived notions, χ^2 (2) = 21.10, $p < .001$, in the post-letters. The difference in the number of reasons between pre- and post-letters was marginally significant on public policy, χ^2 (2) = 5.29, $p = .07$, and not significant on two other domains: environment, χ^2 (2) = 3.79, $p = .15$, and economy, χ^2 (2) = 3.21, $p = .21$.

Another interesting finding is that fewer students included preconceived notions in the post letters. Specifically, the number of letters containing no preconceived notion increased almost three times from pre- (24.6%) to post-letters (64.6%); the number of letters presenting at least one preconceived notion dropped by half from pre- (75.4%) to post-letters (35.4%). These results suggest that students acquired domain knowledge to support their claim and relied less on their personal ideas in decision-making after the Space Exploration Unit. The terms "preconceived notions" or "personal ideas" refer to pre-existing misconceptions or personal beliefs or assumptions not grounded in scientific evidence. Further, the total number of students who integrated at least two domain reasons in the decision letters almost tripled from pre-letters ($n = 7$) to post-letters ($n = 19$).

Table 5. Descriptive Statistics of Length Measures for Decision Letters and Antarctic Essays

	Space Exploration Decision Letters				Antarctic Transfer Essays			
	Treatment Group Pre-Letter (n = 65)		Treatment Group Post-letter (n = 65)		Treatment Group Post-Essay (*n* = 70)		Comparison Group Post-Essay (n = 52)	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Total number of words***	60.25	38.38	97.06	50.91	106.43	47.69	73.29	40.9
Total number of idea units***	5.56	3.48	9.12	5.62	8.27	3.41	5.42	2.65

Note. *** p <.001; The frequency means the number of letters coded

Table 6. Frequencies and Percentages of Decision Letters Providing Domain Reasoning and Preconceived Notions

		Domain Reason Total (%) **	Technology (%) **	Public Policy (%)	Environment (%)	Economy (%)	Preconceived Notion (%) ***
Pre-Letter	No Reason	42(64.6)	55(84.6)	52(80)	58(89.2)	63(96.9)	16(24.6)
(n=65)	One Reason	13(20)	9(13.8)	9(13.8)	4(6.2)	2(3.1)	25(38.5)
	>One Reason	10(15.4)	1(1.5)	4(6.2)	3(4.6)	0	24(36.9)
Post-Letter	No Reason	17(26.)	39(60)	41(63.1)	52(80)	59(90.8)	42(64.6)
(n=65)	One Reason	21(32.3)	22(33.8)	13(20)	11(16.9)	1(1.5)	11(16.9)
	>One Reason	27(41.5)	4(6.1)	11(17)	2(3.1)	5(7.6)	12(18.5)

Note. The total number of domain reasons ranged from 0-6 for pre-letters and from 0-8 for post-letters;
The total number of preconceived notions ranged from 0-7 for pre-letters and from 0-8 for post-letters.
** $p < .01$; *** p <.001 based on Chi-square tests

Argument Type and Quality

Table 7 presents the frequencies and percentages of decision letters coded as one of the four argument types (no argument, one-sided only, dual perspective, and integrative perspective), elaborated reasons, and multi-link reasoning. The Chi-square test showed no significant differences in the distribution of four argument types between the pre- and post-conditions, $\chi^2(3) = 1.49$, $p = .68$. The comparison between the pre- and post-letters on the indicators of the argument quality showed that post-letters had more elaborated reasons (23%) than the pre-letters (9.2%), $\chi^2(1) = 6.12$, $p = .01$; used more evidence (9.2%) than the pre-letters (1.5%), $\chi^2(1) = 4.45$, $p = .03$; and contained more multi-link reasoning (7.7%) than pre-letters (0%), $\chi^2(1) = 5.44$, $p = .02$. Here is an excerpt of a post-intervention decision letter, written by a Spanish-English bilingual girl, that illustrates complex reasoning and integrative perspectives of the pros and cons of increasing funding.

> I think they should decrease because if they increase many people can lose their job because of the taxes... The effect of increasing or decreasing space exploration funding on the economy is not a simple business, jobs and government income will be affected differently if space exploration funding is either increased or decreased it is time for exploration. If they increase it people can lose jobs and not some, a lot of people can lose it and if they decrease they can get more taxes and more people can get more jobs.

RQ2: Treatment Group's Argument Transfer to Post-Antarctic Essays

To answer RQ2: To what extent do treatment students transfer the argumentation skills acquired during the DISCUSS intervention to solve a novel socioscientific issue, and demonstrate improved outcomes when compared with a comparison group?, we first compared the treatment group's post-letters and post-Antarctic essays and then compared treatment and comparison group differences in post-Antarctic essays on each indicator.

Writing length

Table 5 presents the total words and total idea units of space exploration post-decision letters written by the treatment group and post-Antarctic essays written by the treatment and comparison groups. The essay length between the treatment group's decision letters and Antarctic essays was not directly comparable because of different prompts. As for the post-Antarctic essays written by the treatment and comparison groups, Independent-Samples Mann-Whitney U Tests showed that students in the treatment group on average wrote significantly longer ($M = 106.43$, $SD = 47.7$) than those in the comparison group ($M = 73.29$, $SD = 40.90$), Mann-Whitney U =2,628, $p < .001$. Similarly, students in the treatment group ($M = 8.27$, $SD = 3.41$) wrote significantly more idea units than those in the comparison group ($M = 5.42$, $SD = 2.65$), Mann-Whitney U =2809, $p < .001$.

Table 7. Frequency Distribution of Argument Type and Quality Indicators in Treatment Group's Pre-Post Decision Letters and Antarctic Essays by Treatment and Comparison Groups

	Space Exploration Decision Letters					Antarctic Transfer Essays					
	Treatment Group (n=65)	Pre-Letter	Treatment Group (n=65)	Post-Letter	T_Pre vs. T_Post Letter	Treatment Group Post-Essay (n=70)		Comparison Group Post-Essay (n=59)		T_Post Letter vs. T-Post Essay	T Post-Essay vs. C Post-Essay
	N	Percent	N	Percent	P value	N	Percent	N	Percent	P value	P value
Argument Type					*p>.05*					**p <.01*	*p >.05*
No Argument	10	15.40%	12	18.50%		4	5.70%	8	13.60%		
One-sided Only	46	70.70%	41	63.10%		44	62.90%	39	66.10%		
Dual Perspective	9	13.80%	8	12.30%		17	24.30%	9	15.30%		
Integrative Perspective	1	1.50%	3	4.60%		1	1.40%	0	0%		
Elaborated Reason	6	9.20%	15	23%	**p <.05*	10	14.30%	4	6.80%	*p >.05*	**p<.01*
Evidence	1	1.50%	6	9.20%	**p <.05*	24	34.30%	17	28.80%	*p >.05*	**p<.01*
Multi-link Reasoning	0	0.00%	5	7.70%	**p <.05*	13	18.50%	2	3.40%	*p >.05*	**p<.001*
Reasoning					*p >.05*					**p< .001*	**p <.01*
No Reason	42	64.60%	17	26.20%		5	7.10%	10	18.90%		
One Reason	13	20.00%	21	32.30%		3	4.30%	11	20.80%		
More Than One Reason	10	15.40%	27	41.50%		62	88.70%	32	60.30%		

Note. T-treatment, C-comparison

Table 8. Quantity and Quality Measures of Target Group Students' Decision Letters and Transfer Essays

Student	Teacher/Class	Gender	Ethnicity	TOSWRF PR	Pre Letter TW	Post Letter TW	Pre Letter Score	Post Letter Score	Source integration	Antarctic Essay
Zain	Ms. D /GE	M	Asian	68	61	64	2	6	Yes	5
Lylah		F	AA	81	137	204	2	11	Yes	10
Adriel		M	Hispanic	30	51	122	3	3	No	6
Quyen		F	Asian	50	47	88	4	5	Yes	N/A
Antez		M	AA	27	120	56	2	4	Yes	4
Avg.					83.2	106.8	2.6	5.8	4/5	6.25
Domingo	Ms. J /GE	M	Hispanic	N/A	81	N/A	4	N/A	N/A	9
David		M	AA	70	74	33	2	2	No	7
Xavion		M	AA	82	64	82	2	6	Yes	8
Rafa		F	Arabic	N/A	48	159	3	5	No	10
Avg.					66.75	91.33	2.75	4.33	1/3	8.5
Emanuel	Ms. L /BE	M	Hispanic	19	26	71	2	3	No	5
Juanito		M	Hispanic	1	27	166	2	2	No	5
Aranza		F	Hispanic	37	51	93	2	N/A	N/A	8
Maria		F	Hispanic	25	33	99	2	3	No	6
Avg.					34.25	107.25	2	2	0/4	6

Note. TOSWRF PR- Test of Silent Word Reading Fluency Percentile Rank; TW-Total Words, AA- African American, Other- Arabic; GE- General Education, BE- Bilingual Education

Ms. D- African American, Ms. J-African American, Ms. L- Hispanic

Domain Reasons

Table 7 presents the frequencies and percentages of students that had none, one or more than one reason in decision letters and Antarctic essays. Two-by-three cross-tab Chi-square analyses showed that the difference in the reason distribution (no reason, one reason, and more than one reason) was significant between the treatment group's post-letters and post-essays, χ^2 (2) = 49.58, $p < .001$. Nearly half of the post-letters (42%) contained more than one reason, and double post-essays (89%) had more than one reason.

Two-by-three cross-tab Chi-square analyses showed the reason distribution was significantly different between the treatment and comparison groups' post-essays, χ^2 (2) = 13.72, $p = .001$. Compared to the essays written by the comparison group (60.3%), 88.7% of the essays written by the treatment group students contained more than one reason to support their claim. More students in the comparison group (18.9%) had no reason than the treatment group (7.1%). Also, more students in the comparison group (20.8%) had only one reason than the treatment group (4.3%). These results support more complex reasoning in the essays written by the treatment group than in the essays written by the comparison group.

Argument Type and Quality

Table 7 presents the frequencies and percentages of post-letters and post-Antarctic essays coded as argument type (no argument, one-sided only, dual perspectives, and integrative perspective) and each argument quality indicator. Two-by-four crosstab Chi-square test showed a significant difference in the distribution of four argument types between the treatment group's post-letters vs. post-essays, $\chi^2(3)$ = 13.30, $p = .004$. It appears that most post-decision letters were single-sided (63.10%) and similarly most post-essays were one-sided (62.9%), however, post-essays written by the treatment group presented almost double dual perspectives (24.3%) than the post-letters (12.3%), suggesting the positive transfer of written argument skills.

The Chi-square test showed no significant difference between the treatment vs. comparison groups on the distribution of four argument types in post-essays, $\chi^2(3)$ = 4.30, $p = .23$. Although not statistically significant, more post-essays written by treatment classes considered both sides of the issues (dual perspectives) (24.3%) than the post-essays written by the comparison classes (14.3%). Few essays demonstrated integrative perspectives. The difference between the treatment and comparison essays on other indicators of the argument quality showed that the essays written by the treatment group had more elaborated reasons (14.3% vs. 6.8%), used more evidence (34.3% vs. 28.8%), and contained more multi-link reasoning (18.5% vs. 3.4%) than the essays written by the comparison group, $\chi^2(1)$ = 8.06, $p = .005$.

The essay below, written by an Asian male and bilingual student, illustrates elaborated reasons supporting increased funding for Antarctic research and multi-link reasoning about the consequence of global warming and Antarctic ice sheet melting.

> We should fund research for Antarctica cause, Antarctica ice sheets are melting rapidly at a hundred cubic kilometers per year. The massive loss of the ice sheet raises the sea level. When the sea level rises it causes huge flooding. If we don't try to find a solution for global warming and continue with our daily lives, the melting rate in Antarctica will increase. The faster the ice sheets melt, the faster the sea level will rise, causing flooding problems all over the world. The flooding would cause chaos and war for land. If we don't do something about global warming now, history will soon repeat itself.

RQ3: Target Students' Writing and Discussions

To examine the connection between student talk and writing outcomes, we analyzed three target groups' writing and classroom talk throughout the DISCUSS intervention. Each target group was a representative cross-section of the treatment class. Table 8 presents the demographic characteristics, English reading fluency of the students from three target groups, and their argumentative writing features.

Reasoning in Writing

Student letter/essay lengths and holistic written argument quality varied across groups. The target groups in the two general education classes showed greater improvement in decision letter holistic scores than the target group in the bilingual education class. Student writing quality in Ms. L's bilingual science classroom was generally low (2 or 3 holistic scores), showed little or no improvement, and no integration of multiple sources. Among the 13 target students, 8 students showed increasing holistic scores. Of those 6 were from the general education classes. To illustrate the pre-post improvement of decision letters, two students' letters (one from the general education class and one from the bilingual education class) are shown below. All student names were pseudonyms.

The letters below were written by Zain, an Asian bilingual boy in Ms. D's general education class (TOSWRF PR 68). Although the letters were roughly the same length, the post-letter showed greater reasoning than the pre-letter. The student presented some preconceived notions in the pre-letter, needing more money to build a space shuttle and purchase equipment for astronauts. In the post-letter, the student provided reasons for increased funding from the perspective of medical technology innovation and presented scientists' views and military officials' views about space exploration. Most excitingly, the student considered alternative perspectives in the end 'although I can see why people disagree.'

Pre-Decision Letter

> I would go with increase because they can use it for more useful materials for one-man ships, more advanced suits for astronauts. It's also because many other astronauts can go on more than 1 mission. The result of that is that NASA and other space organizations can spread, so now we can do more than one mission.

Post-Decision Letter

> I recommend space exploration funding should be increased because space exploration has taken us to many advantages. I also think that it can help us find cures for people who have a bad disease. I think the military needs an increase so when it comes to war they have the upper hand, although I can see why people disagree.

The following Antarctic essay written by Aranza, a Spanish-English bilingual girl in Ms. L's bilingual education class (TOSWRF PR 37), presented a clear claim and textual evidence to support funding research in Antarctica. Although the essay remained one-sided, she provided multi-link reasoning about the chains of consequences of melting ice sheets in Antarctica.

> The United States has spent approximately $300 million for every year to research in Antarctica. Yes we should fund research in Antarctica. Yes, we should fund research in Antarctica or maybe people that go to Antarctica have to tell us if bad things are happening there. Data from space satellites show that Antarctica has been losing more than one hundred cubic kilometers of ice each year since 2002. If the Antarctica ice sheet continues to melt, global sea levels will rise which can result in the flooding of coastal cities like New York, Miami, and Galveston.

Source Integration

Five out of 13 students integrated at least two sources of reasons from four domains (technology, economy, environment and public policy) in their post-intervention decision letters. However, source integration was not observed in the bilingual science classroom.

Student Talk

To further illustrate target group students' participation in small-group discussions, Table 9 provides student talk coding results. Student talk patterns varied by group. Overall, students talked more in the bilingual education class than in the two general education classes. However, student CER talk was far more evident (Ms. J: 26% and Ms. D: 18%), and student questions were more prevalent (Ms. J: 5% and Ms. D: 6%) in the two general education classes than in the bilingual education class

Table 9. Percentages of Target Group Student Discussion Codes (N=13)

Student	Teacher/Class	Claim	Evidence	Reasoning	Question	Rebuttal	Nontransactive	Cumulative	Exploratory	Total Turn
Zain	Ms. D /GE	18.59%	4.17%	15.64%	0.00%	0.00%	65.03%	22.10%	12.87%	15
Lylah		5.88%	1.96%	12.25%	33.33%	0.00%	75.98%	14.22%	9.80%	7
Adriel		8.33%	0.00%	16.67%	0.00%	0.00%	79.17%	20.83%	0.00%	7
Quyen		0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0
Antez		0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0
	Avg.	6.56%	1.23%	8.91%	6.67%	0.00%	44.04%	11.43%	4.53%	5.8
Domingo	Ms. J /GE	17.78%	0.00%	18.89%	6.67%	0.00%	60.00%	34.44%	5.56%	7
David		20.74%	6.67%	20.74%	10.37%	3.70%	40.85%	45.84%	13.30%	13
Xavion		5.33%	0.00%	9.75%	3.93%	0.00%	45.98%	44.40%	9.62%	25
Rafa		7.50%	0.00%	3.33%	0.00%	0.00%	21.48%	38.52%	6.67%	6
	Avg.	12.84%	1.67%	13.18%	5.24%	0.93%	42.08%	40.80%	8.79%	12.75
Emanuel	Ms. L /BE	4.27%	0.00%	0.85%	1.71%	0.85%	98.29%	0.00%	1.71%	20
Juanito		5.56%	0.00%	1.85%	3.70%	3.70%	97.35%	0.00%	2.65%	23
Aranza		2.25%	0.90%	2.25%	1.35%	0.00%	94.62%	3.12%	2.25%	40
Maria		1.33%	0.00%	0.67%	4.67%	0.67%	92.29%	4.17%	3.54%	32
Avg.	Avg.	3.35%	0.23%	1.41%	2.86%	1.31%	95.64%	1.82%	2.54%	28.75

Note. The percentages averaged three videotaped lessons; GE- General Education, BE-Bilingual Education

Table 10. Pearson Correlation between Discussion Indicators and Writing Measures (N=13)

	Pre_letter TW	Post_letter TW	Pre_letter Holistic Scores	Post_letter Holistic Scores	Pre-Post Score Growth	Source Integration	Antarctic _Essay
Claim	0.107	-0.253	0.096	0.037	0.071	-0.091	0.156
Evidence	0.236	-0.369	-0.326	0.043	0.112	0.035	-0.009
Reasoning	0.328	-0.111	0.139	0.23	0.251	0.062	0.283
Question	0.643*	0.513	-0.201	0.559*	0.736**	0.236	0.512
Rebuttal	-0.263	-0.072	-0.323	-0.411	-0.37	-0.465	-0.281
Nontransactive	-0.373	0.274	-0.411	-0.236	-0.087	-0.523	-0.104
Cumulative	0.161	-0.107	0.138	0.198	0.189	0.002	0.548
Exploratory	0.288	-0.072	-0.303	0.412	0.499	0.221	0.371
Total_Turn	-0.486	-0.082	-0.548	-0.482	-0.3	-0.437	-0.076

Note. TW- total words

(CER: 5% and question: 2% in Ms. L's class). Nontrasactive talk took up about 40% of all student talk, and cumulative talk and exploratory talk accounted for 13% and 15% in the two general education classes. Over 95% of student talk in the bilingual science classroom was nontransactive and independent, suggesting the absence of dialogic interaction.

Nontransactive talk

Most of the classroom talk in the bilingual science classroom was clarifying vocabulary and teacher questions were low-level. Although the teacher constantly promoted students to talk with a neighbor student or among the table mates, students mainly answered teacher questions with short responses most often in Spanish. The teacher recognized students' struggle with academic vocabulary, reading, and making a connection with science content and encouraged students to draw the connection between their home language and English academic vocabulary. The excerpt below from Ms. L's bilingual class illustrates the home language and academic vocabulary connection.

> Teacher: Look at the front. Tell them, what do you think about when you say tele? Telephone, Television, so what do you think of telemedicine?
>
> Aranza: Telemedicina…Estan en la tele? [Repeats telemedicine in Spanish]. Are they on TV?
>
> Teacher: Yes! It's a virtual doctor ... he is a real doctor but you'll see him on the internet or the phone. You don't have to go to his office in person….

Cumulative talk

The excerpt below from the target group in Ms. D's general education class illustrates cumulative talk. Students were working collaboratively to create a Venn diagram of the inner and outer planets. Students were building on one another and the peer responses were confirmative and expanding.

> Teacher: We're getting ready to quickly do a Venn diagram from the inner and outer planets. So everybody should have this. Everybody has an inner and outer planet Venn Diagram. Okay, so, we are going to — you're going to get your textbooks in a minute. You're going to discuss at your tables for about thirty seconds. What do you know about those inner and outer planets, anything you might know about them? Go.
>
> Lylah: Inner planets kind of sound like...
>
> Zain: I have no idea.
>
> Lylah: I think the inner planets are closer to the sun, and the outer planets are further.

Zain: I think the inner planets would be the first four.

Adriel: And outer planets are the last four. And inner planets are hotter than the outer planets.

Zain: Yeah, because it's closer to the sun.

Adriel: And outer planets are colder. And the inner planets um... the inner planets...

Linda: She said to write inside it?

Adriel: There's eight planets, right?

Student: Six.

Zain: No, there's eight.

Adriel: Eight? What is it? Mercury, Venus, Mars?

Zain: Mercury, Venus, Earth, Mars, Jupiter, Saturn, Uranus, Neptune...

Exploratory talk

The exploratory talk was rare, less than 10% of talk with a slightly higher percentage in Ms. J's general education class. The excerpt below from Ms. J's target group illustrates exploratory talk. During the lesson, students read and discussed the impact of space exploration on technological innovation. Before the excerpt, students looked for claims and evidence about the positive impact of space exploration on medical technology. They respectfully challenged one another and the disagreement was justified.

David: See it would, it would increase the funding for space exploration.

Xavion: (cutting David off) ... innovation one, okay let's think about all technology. medical, environmental, you know all that and we like just decreased it then we will have fewer MRIs, less money to see storms that are coming out and what's going on in the world for global warming and everything like that.

David: I say that um increasing funding for space exploration would be good for technological innovation because like since we're going to like who knows if we are or not but if we explore more planets and other moons, then we're going to need to build more technological innovation because like um, we can't use the same technology to land on the moon. Like we can't use that same technology if we're going to land on Mars.

Xavion: Because first of all, the rockets barely had enough fuel to go to the moon probably, if we try to go further away, feel bad for the men! Men and women are in space because they're going to be stranded in space.

Rafa: I think that they should increase it because you know when um in outer space, they're going to explore more things and the more stuff they explore the more people will get interested, and the more people are interested the more people work there and the more people that work there, the more the people that will help to build the things and if you explore enough and find a planet, maybe the space trash we can land it there instead of polluting it on earth.

David: (Excitingly) Oh Yeah! Space Trash! Yeah since we can like clean the earth.

The connection between Student Talk and Writing

Table 10 presents the Pearson correlations between the student talk indicators and writing quantity and quality measures. Given the small sample size, the correlations were mostly positive but relatively low in general. However, several interesting patterns emerged. First, student questions were positively and significantly correlated with student writing length and quality measures. Exploratory talk was also positively correlated with post-decision letter holistic scores and transfer essay quality. However, nontransactive talk was negatively correlated with writing outcomes. These results suggest the promise of dialogic inquiry and argumentation in promoting student writing.

Students who demonstrated improved writing quality from pre- to post-letters in two general education classes (Zain, Lylah, Xavion, and Rafa) also talked more and engaged more in the dialogic talk — cumulative talk and exploratory talk. Three of these students (Zain, Lylah, Xavion) utilized source integration in their post-decision letters. Students' capacity to integrate resources from various domains illustrates better reasoning ability (Barzilai et al., 2015). These results highlight the connection between student talk patterns and their writing.

8. Discussion

The current study examined the impact of a literacy-infused and SSI-based approach in science classrooms on the argumentative writing of linguistically diverse sixth-grade students in an urban intermediate school. The current results suggest the promise of SSI-based instruction and literacy integration in enhancing diverse students' argumentative thinking and writing in science classrooms. The findings have extended the literature on dialogic inquiry and argumentative writing in several ways. First, the DISCUSS curriculum enhances students' ability to make informed decisions by drawing on several sources of evidence to support their claims. In the decision letters, despite the disproportional distribution of reasoning sources from the four domains, with most reasons from the technology perspectives and least reasons representing economic perspectives, the results demonstrate that after participating in the DISCUSS curriculum, students were more able to draw on varied sources of evidence in decision making. Similar to Barzilai et al.'s (2015) detailed account of college students' spontaneous sourcing practices as they engage in reading divergent expert accounts about socioscientific issues, the current results suggested that better capacity to draw evidence from various sources were related to more complex argumentation and increased claim justification in participants' arguments.

Second, compared to the pre-letters, students were more able to use evidence-based reasoning and relied less on personal experiences and beliefs in the post-letters. The results showed gains in presenting elaborated reasons, evidence, and multi-link causal reasoning in the post-letters than in the pre-letters. This study provides evidence about how SSI-based curriculum and argumentative discourse can promote early adolescents' evidence-based reasoning. This finding is significant because middle school students often construct explanations based on personal ideas instead of explanations from evidence (Driver et al., 1996; Duhalongsod, 2017). Previous research shows that middle and high school students often lack specific evidence and explanation to engage in academic discussions or debates. For example, Walker and Zeidler (2007) found that high school students often use personal attacks and fallacious argumentation (e.g., extreme examples, erroneous grounds) during debates about socioscientific issues. The major barrier is the lack of background knowledge and argumentation skills to engage in meaningful and productive debates. Duhalongsod (2017) documented middle school students' progress from personal attacks to evidence and reasoning after participating in the Social Studies Generation program, which aims to promote students' discussion and argumentation around controversial public issues through classroom debates.

The current findings suggest that students were more able to use claim-evidence-reasoning in their argumentative writing from the pre- to post-decision letters. Berland and Reiser (2009) identified two ways students connect the evidence to the claim in science argument: (a) include background knowledge and (b) describe the logical connection between the evidence and claim. The current findings showed that students heavily relied on their prior knowledge and simple logic to make decisions before the Space Exploration Unit, however, post-letters contained more sophisticated logic and integrated multiple sources of evidence to make connections between claims and evidence.

Third, students who participated in the DISCUSS intervention were able to apply these acquired argument skills to reason about a novel SSI: *Should we fund* *research in Antarctica?* Students in the treatment group who were more likely to draw multiple reasons to support their claims, elaborate their reasons, and use evidence and multi-link reasoning in the post-decision letters presented higher-quality written arguments in the Antarctic transfer essays. These results provide evidence of the argument transfer across different SSIs for the treatment group.

Regarding the plausible explanations underlying the transfer of argument skills in existing literature, we attribute such transfer to change in thinking dispositions (Zohar & Nemet, 2002), metacognition about the well-constructed argument (Koedinger & Wiese, 2015; Reznitskaya et al., 2008), and expansive framing contexts (Engle, 2011; 2012) that map onto our design of DISCUSS intervention. First, students who participated in the DISCUSS curriculum have the opportunity to become active agents who engage in role-playing and chains of reasoning, compared to the no-treatment students where teachers do most of the questioning and evaluating, leaving students with circumscribed opportunities for extended reasoning, independent thinking and decision making. The change of classroom participation norm from the traditional IRE (Initiate-Response-Evaluate) pattern to open participation in classroom dialogues in the DISCUSS treatment condition shapes what students think is expected of them and thus changes their patterns of responses. There are multiple reasons why formally guiding students in interactive argumentation for academic contexts is necessary (Noroozi et al., 2012). Students who are set in their beliefs may be reluctant to accept opposing viewpoints that are incompatible with their own. They may also feel that counterarguments from others are personal attacks. These initial student behaviors can be exacerbated in the case of emergent bilinguals as they develop English language proficiency and content knowledge simultaneously. The DISCUSS curriculum was designed to provide academic and language scaffolding for emergent bilinguals and to promote their science content knowledge development.

Second, the DISCUSS curriculum afforded multiple opportunities through reading, discussion, and writing for students to learn about and practice making decisions about complex and controversial issues, which allows students to acquire an abstract meta-level awareness of argumentation and scientific discourse, called argument schema (Reznitskaya et al., 2008), and transfer such knowledge to a novel task. Previous research has provided theoretical and empirical support for the argument schema which enables the transfer of argument skills from oral group discussions to individual writing (Reznitskaya et al., 2001).

Third, DISCUSS curriculum features expansive framing which may have promoted the transfer of argument skills (Engle et al., 2012). At the beginning of DISCUSS lessons, students role-played the Space Council members and were charged with making an important decision regarding the space budget. Students reviewed the newsletter that presented the historical data on the space budget and engaged in a larger societal conversation about the implications of the space budget cut. The expansive framing of the SSI was extended across times, locations, and learning activities in the DISCUSS curriculum. However, in the comparison group, the space science unit was delivered in a bounded framing where students were not afforded agency to solve a greater problem.

Although the current results from both the decision letters and Antarctic transfer essays show positive effects on students' evidence-based reasoning in argumentative writing, students are generally less able to incorporate the counter-argument and weigh the importance of reasons or opposing perspectives to reevaluate their claims. Most arguments in both writing tasks were one-sided or myside biased, meaning arguments citing positive attributes of the favored position (Kuhn & Crowell, 2011). However, the Antarctic essays written by treatment classes were almost twice more likely to consider both sides of the issues (dual perspectives) than the essays written by the comparison classes. The difference in disposition to consider both sides of issues could be traced to fundamental differences in classroom dialogue between the DISCUSS and traditional science classrooms. Although the current study does not provide direct evidence about the difference in the discourse patterns in the treatment and comparison classes, our classroom observations suggest that the DISCUSS classes emphasized more NGSS-aligned science practices such as Claim-Evidence-Reasoning-Rebuttal (CERR) and classroom discussions, whereas the comparison science classrooms maintained traditional teacher-centered instruction where students had limited opportunities to ask questions and engage in extended talk. Discourse analyses of the target group discussions revealed that students who wrote improved decision letters from pre- to post-DISCUSS intervention talked more and engaged more in connected or cumulative talk and exploratory talk. As shown in the excerpts of the results, students who wrote better argumentative essays engaged in dialogic argumentation more often, whether the goal was consensus-building (accumulative talk) or persuasion involving confronting opinions different from one's own (exploratory talk).

Nonetheless, in the case of the target group in the bilingual science class, students engage mostly in non-dialogic talk (non-transactive and independent) and the quality of their writing is developing. There are several reasons why students in the bilingual class participated less in classroom discussions and benefited less in writing. First, in an observational study of the bilingual treatment class, we found that the teacher used spontaneous Spanish mainly for classroom management purposes like redirecting behavior and reiterating instructions previously given in English (Enriquez-Andrade, Wui, Zhang, Relyea, & Wong, under review). The teacher was reluctant to release the control and give space for extended dialogue. Intentional use of pedagogical translanguaging by allowing students to use their full linguistic repertoire in classroom talk may enhance student participation and science learning (Licona & Kelly, 2019).

Second, students needed support in reading the texts, understanding the key academic vocabulary, and building background knowledge about the space budget issue. Although peer reading was encouraged in the bilingual class, most of the student talk focused on clarifying vocabulary in Spanish and less on reading comprehension. Third, source-based writing tasks in the current study were challenging to emergent bilinguals because students were expected to identify argument elements in texts, integrate and evaluate information from different sources when writing, and express complex ideas in a persuasive genre. Future work should consider extending instructional support to empower emergent bilingual students to participate in extended dialogues and scaffolded writing in science classrooms. Such instructional supports for emergent bilinguals may include guided prompts for peer talk, teacher explicit instruction and modeling of writing purposes and processes, and the use of graphic organizers to outline ideas and argument elements (Square & Clark, 2020); allowing students to use multimodal representations (visuals, drawings, etc.) and translanguaging in expressing ideas, and engaging students in collaborative writing and revisions (Ardasheva et al., 2015).

There are several limitations in the current study. First, the current study primarily used a mixed-methods, single-group pretest-posttest study combined with a post-hoc comparison with a no-treatment group. The decision letter writing task was administered to the treatment classes only. Because the comparison students did not receive instruction in writing or were not exposed to the argument texts associated with space exploration, it would not be fair to ask them to write a decision letter. Therefore, the single group pre-post comparison for the decision letters may reduce the internal validity of the findings and cannot explain why the instruction was effective or not effective. The post-hoc comparison with a no-treatment group on a near transfer measure provides some evidence of argument transfer of the treatment group, but a more remote transfer task is needed in the future combined with a comparison group. Second, nonsignificant differences between treatment conditions on argument type and quality may be due to the study being underpowered. Due to the small sample size, the data were not disaggregated by participants' language status. Future research should investigate whether similar interventions benefit English-only speakers and emergent bilingual students differently and how the intervention can be adapted to enhance emergent bilingual students' participation and learning. Third, since few essays exhibit the integrative perspectives of a complex argument, student writing in the current study does not reflect the complex argumentation that recognizes tensions when considering two competing options.

The current study makes a significant contribution to the limited theoretical underpinnings and empirical research on writing in science, especially for emergent bilinguals. In particular, the findings enrich our understanding of argumentative writing development and effective approaches to argumentative writing instruction in content areas. Although the disaggregated findings by language status (EBs vs. Non-EBs) were not possible due to the small sample size in this study, the findings suggest the value of a language-rich SSI-based approach to science and language integration to enhance student argumentative writing for all students including emergent bilinguals. These findings have implications for curriculum development, teaching, and assessment strategies to integrate science, language, and literacy instruction for emergent bilingual students. Through carefully designed curriculum and teacher scaffolding of oral and written discourse, emergent bilingual students can be empowered and prepared to engage in deep reasoning regarding critical community and societal issues. Teacher scaffolding strategies in four phases of DISCUSS activities: *framing, reading, discussion,* and *writing* are discussed in more detail by Zhang, Lee, Iluore, Relyea, and Wui (2022).

In the future, it would be interesting to examine how science content knowledge relates to argument writing skills. Much remains to be understood about the role of argumentation and writing in learning science and how to improve science teachers' capacity to facilitate writing and argumentation to learn science, in particular, how to encourage students to incorporate different perspectives by adding self-revision or peer feedback in writing. Future research should investigate the civic or social justice aspects of argumentative writing — whether students employ a public voice or audience awareness and use writing as a tool to promote thinking, learning, and participation in society.

Acknowledgements

This study is supported by a startup grant awarded to the first author by the University of Houston. The authors would like to thank the participating teachers, students, and parents. We appreciate Jennifer Donze and Uchenna Emenaha's assistance in curriculum development, as well as Haelim Jeong, Yumei Li, and Lana Kharabi-Yamato's assistance in transcribing and data preparation. The authors are also indebted to the editor's and reviewers' insightful feedback on the earlier versions of the manuscript.

References

Adey, P., & Shayer, M. (2015). The effects of cognitive acceleration. In L.R. Resnick, C.S.C. Asterhan, & S. N. Clarke (Eds.), *Socializing intelligence through academic talk and dialogue* (pp. 127-140). Washington, DC: American Educational Research Association. https://doi.org/10.3102/978-0-935302-43-1_10

Ardasheva, Y., Norton-Meier, L., & Hand, B. (2015). Negotiation, embeddedness, and non-threatening learning environments as themes of science and language convergence for English language learners. *Studies in Science Education, 51*(2), 201–249. https://doi.org/10.1080/03057267.2015.1078019

Atabey, N., & Topcu, M.S. (2017). The development of a socioscientific issues-based curriculum unit for middle school students: Global warming issue. *International Journal of Education in Mathematics, Science and Technology (IJEMST), 5*(3), 153-170. https://doi.org/10.18404/ijemst.296027

August, D., Branum-Martin, L., Cardenas-Hagan, E., Francis, D. J., Powell, J., Moore, S., & Haynes, E. F. (2014). Helping ELLs meet the common core state standards for literacy in science: The impact of an instructional intervention focused on academic language. *Journal of Research on Educational Effectiveness*, *7*(1), 54–82. https://doi.org/10.1080/19345747.2013.836763

Barzilai, S., Tzadok, E., & Eshet-Alkalai, Y. (2015). Sourcing while reading divergent expert accounts: Pathways from views of knowing to written argumentation. *Instructional Science, 43*, 737–766. https://doi.org/10.1007/s11251-015-9359-4

Belland, B. R., Gu, J., Armbrust, S., & Cook, B. (2015). Scaffolding argumentation about water quality: A mixed-method study in a rural middle school. *Educational Technology Research and Development, 63*(3), 325-353.https://doi.org/10.1007/s11423-015-9373-x

Berland, L., & Reiser, B. (2009). Making sense of argumentation and explanation. *Science Education, 93*, 26-55.https://doi.org/10.1002/sce.20286

Cavagnetto, A. R. (2010). Argument to foster scientific literacy: A review of argument interventions in K–12 science contexts. *Review of Educational Research*, *80*(3), 336–371. https://doi.org/10.3102/0034654310376953

Cervetti, G. N., Barber, J., Dorph, R., Pearson, P. D., & Goldschmidt, P. G. (2012). The impact of an integrated approach to science and literacy in elementary school classrooms. *Journal of Research in Science Teaching, 49*(5), 631–658.https://doi.org/10.1002/tea.21015Chan, C., Burtis, J., & Bereiter, C. (1997). Knowledge building as a mediator of conflict in conceptual change. *Cognition and Instruction, 15*, 1-40.https://doi.org/10.1207/s1532690xci1501_1

Chen, Y.C., Hand, B., & Park, S. (2016). Examining elementary students' development of oral and written argumentation practices through argument-based inquiry. *Science & Education*, *25*, 277–320. https://doi.org/10.1007/s11191-016-9811-0

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

Chowning, J.T., Griswold, J. C., Kovarik, D. N., & Collins, L. J. (2012). Fostering critical thinking, reasoning, and argumentation skills through bioethics education, *Plos One, 7*(5), e36791. https://doi.org/10.1371/journal.pone.0036791

Clark, A. M., Anderson, R. C., Archodidou, A., Nguyen-Jahiel, K., Kuo, L.-J., & Kim, I. (2003). Collaborative reasoning: Expanding ways for children to talk and think in the classroom. *Educational Psychology Review, 15*, 181-198

Crowell, A., & Kuhn, D. (2014). Developing dialogic argumentation skills: A 3-year intervention study. *Journal of Cognition and Development 15*(2), 363–381.

de Oliveira, L. C., & Lan, S. (2014). Writing science in an upper elementary classroom: A genre-based approach to teaching English language learners. *Journal of Second Language Writing, 25*, 23–39. https://doi.org/10.1016/j.jslw.2014.05.001

Dawson, V., & Carson, K. (2020). Introducing argumentation about climate change: Socioscientific issues in a disadvantaged school. *Research in Science Education 50*, 863–883. https://doi.org/10.1007/s11165-018-9715-x

Doise, W., & Mugny, G. (1984). *The Social development of the intellect*. Oxford: Pergamon Press (originally published, 1981).

Dolan, T., Nichols, B., & Zeidler, D. (2009). Using socioscientific issues in primary classrooms. *Journal of Elementary Science Education, 21*(3), 1-12. https://doi.org/10.1007/BF03174719

Driver, R., Leach, J., Millar R., & Scott P. (1996). *Young people's images of science*. Bristol, PA: Open University Press.

Duhalongsod, L. (2017). Classroom debates in middle school social studies: Moving from personal attacks to evidence and reasoning. *Middle Grades Research Journal, 11*(2), 99-115.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York: Random House.

Engle, R. A., Nguyen, P. D., & Mendelson, A. (2011). The influence of framing on transfer: Initial evidence from a tutoring experiment. *Instructional Science, 39*(1), 603–628. https://doi.org/10.1007/s11251-010-9145-2

Engle, R. A., Lam, D. P., Meyer, X. S., & Nix, S. E. (2012). How does expansive framing promote transfer? Several proposed explanations and a research agenda for investigating them. *Educational Psychologist, 47*(3), 215-231. https://doi.org/10.1080/00461520.2012.695678

Enriquez-Andrade, A., Wui, M.G., Zhang, J., Relyea, J., & Wong, S. (under review). Teachers' Language Ideologies and Practices on the Use of Spanish in Middle School Science Classrooms.

Foong, C., & Daniel, E.G.S. (2013). Students' argumentation skills across two socio- scientific issues in a Confucian classroom: Is transfer possible?. *International Journal of Science Education, 35*(14), 2331-2355. https://doi.org/10.1080/09500693.2012.697209

Garza, T., Huerta, M., Lara-Alecio, R., Irby, B. J., & Tong F. (2018). Pedagogical differences during a science and language intervention for English language learners. *The Journal of Educational Research, 111*(4), 487-496. https://doi.org 10.1080/00220671.2017.1302913

George, R. (2000). Measuring change in students' attitudes toward science over time: An application of latent variable growth modeling. *Journal of Science Education and Technology, 9*(3), 213-225.

González-Howard, M., & McNeill, K.L. (2019). Teachers' framing of argumentation goals: Working together to develop individual versus communal understanding. *Journal of Research in Science Teaching, 56*(6), 821-844. https://doi.org/10.1002/tea.21530

González-Howard, M., McNeill, K. L., Marco-Bujosa, L. M., & Proctor, C. P. (2017). 'Does it answer the question or is it French fries?': an exploration of language supports for scientific argumentation. *International Journal of Science Education, 39*(5), 528-547. https://doi.org/10.1080/09500693.2017.1294785

Gutierez, S. B. (2015). Integrating socio-scientific issues to enhance the bioethical decision-making skills of high school students. *International Education Studies, 8*(1), 142-151. https://doi.org/10.5539/ies.v8n1p142

Hand, B., Norton-Meier, L. A., Gunel, M., & Akkus, R. (2016). Aligning teaching to learning: a 3-year study examining the embedding of language and argumentation into elementary science classrooms. *International Journal of Science and Mathematics Education, 14*(5), 847–863. https://doi.org/10.1007/s10763-015-9622-9

Herman, B. C., Newton, M. H., & Zeidler, D. L. (2021). Impact of place-based socioscientific issues instruction on students' contextualization of socioscientific orientations. *Science Education*, *105*, 585– 627. https://doi.org/10.1002/sce.21618

Huerta, M., & Garza, T. (2019). Writing in science: Why, how, and for whom?: A systematic literature review of 20 years of intervention research (1996–2016). *Educational Psychological Review, 31*, 533–570. https://doi.org/10.1007/s10648-019-09477-1

Iordanou, K. (2010). Developing argumentative skills across scientific and social domains. *Journal of Cognition and Development, 11*(3), 293-327. https://doi.org/10.1080/15248372.2010.485335

Khishfe, R. (2014). Explicit nature of science and argumentation instruction in the context of socioscientific Issues: An effect on student learning and transfer. *International Journal of Science Education, 36* (6), 974-1016. https://doi.org/ 10.1080/09500693.2013.832004

Khishfe, R. (2013). Transfer of nature of science understandings into similar contexts: Promises and possibilities of an explicit reflective approach. *International Journal of Science Education, 35*(17), 2928-2953. https://doi.org/10.1080/09500693.2012.672774

Kim, S. L., & Kim, D. (2021). English learners' science-literacy practice through explicit writing instruction in invention-based learning. *International Journal of Educational Research Open, 2,* 100029. https://doi.org/10.1016/j.ijedro.2020.100029

Koedinger, K.R., & Wiese, E. S. (2015). Accounting for socializing intelligence with the knowledge-learning-instruction framework. In L.B. Resnick, C. Asterhan, & S.N. Clarke (Eds.), *Socializing intelligence through academic talk and dialogue* (pp. 275-286). Washington, DC: American Educational Research Association. https://doi.org/10.3102/978-0-935302-43-1_22

Kuhn, D. (1991). *The skills of argument*. New York, NY: Press Syndicate of the University of Cambridge.

Kuhn, D. (2015). Thinking together and alone. *Educational Researcher*, *44*(1), 46–53. https://doi.org/ 10.3102/0013189X15569530

Kuhn, D., & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychological Science, 22*(4), 545-552. https://doi.org/10.1177/0956797611402512

Kuhn, D., Hemberger, L., & Khait, V. (2016). Tracing the development of argumentive writing in a discourse-rich context. *Written Communication*, *33*(1), 92–121. https://doi.org/10.1177/0741088315617157

Kuhn, D., Zillmer, N., Crowell, A., & Zavala, J. (2013). Developing norms of argumentation: Metacognitive, epistemological, and social dimensions of developing argumentative competence. *Cognition and Instruction, 31*(4), 456-496. https://doi.org/10.1080/07370008.2013.830618

Lara-Alecio, R., Tong, F., Irby, B. J., Guerrero, C., Huerta, M., & Fan, Y. (2012). The effect of an instructional intervention on middle school learners' science and English reading achievement. *Journal of Research in Science Teaching, 49*(8), 987–1011. https://doi.org/10.1002/tea.21031.

Lee, O., Mahotiere, M., Salinas, A., Penfield, R. D., & Maerten-Rivera, J. (2009). Science writing achievement among English language learners: Results of three-year intervention in urban elementary schools. *Bilingual Research Journal, 32*(2), 153-167. https://doi.org/10.1080/15235880903170009

Lee, O., Penfield, R. D., & Buxton, C. (2011). Relationship between "form" and "content" in science writing among English language learners. *Teachers College Record, 113*(7), 1401–1434.

Licona, P. R., & Kelly, G. J. (2019). Translanguaging in a middle school science classroom: Constructing scientific arguments in English and Spanish. *Cultural Studies of Science Education, 15*, 485–510. https://doi.org/10.1007/s11422-019-09946-7

Mather, N., Hammill, D. D., Allen, E. A., & Roberts, R. (2004). *Test of silent word reading fluency*. Austin, TX: Pro-Ed.

Mayer, R. E. (1985). Structural analysis of science prose: Can we increase problem solving performance? In B. K. Britton & J. B. Black (Eds.), *Understanding of expository text* (pp. 65-87). Hillsdale, NJ: Erlbaum. https://doi.org/10.4324/9781315099958-3

McNeill, K. L., & Krajcik, J. (2008). Scientific explanations: Characterizing and evaluating the effects of teachers' instructional practices on student learning. *Journal of Research in Science Teaching, 45*(1), 53–78.

McNeill, K.L., & J. Krajcik, J. (2011). *Supporting grade 5–8 students in constructing explanations in science: The claim, evidence and reasoning framework for talk and writing*. New York: Pearson, Allyn, and Bacon.

McNeill, K.L., & Krajcik, J. (2012). *Supporting grade 5–8 students in constructing explanations in science: The claim, evidence, and reasoning framework for talk and writing*. Upper Saddle River, NJ: Pearson Education.

McNeill, K. L., & Pimentel, D. (2010). Scientific discourse in three urban classrooms: The role of the teacher in engaging high school students in argumentation. *Science Education, 94*(2), 203-229. https://doi.org/10.1002/sce.20364

Mercer, N. (1996). The quality of talk in children's collaborative activity in the class room. *Learning and Instruction, 6*, 359-378.

Morris, J. A., Miller, B. W., Anderson, R. C., Nguyen-Jahiel, K. T., Lin, T.-J., Scott, T., . . . Ma, S. (2018). Instructional discourse and argumentative writing. *International Journal of Educational Research,* 90(1), 234-247. https://doi.org/10.1016/j.ijer.2018.03.001

Murphy P. K., Greene J. A., Allen, E., Baszczewski, S., Swearingen, A., Wei, L., & Butler, A. M. (2018). Fostering high school students' conceptual understanding and argumentation performance in science through Quality Talk discussions, *Science Education, 102*, 1239-1264. https://doi.org/10.1002/sce.21471

National Assessment of Educational Progress (NAEP). (2011). *Writing 2011.* Retrieved from https://nces.ed.gov/nationsreportcard/pdf/main2011/2012470.pdf

National Research Council. 2015. *Guide to implementing the Next Generation Science Standards.* Washington, DC: The National Academies Press. https://doi.org/10.17226/18802.

Newell, G. E., Beach, R., Smith, J., & VanDerHeide, J. (2011). Teaching and learning argumentative reading and writing: A review of research. *Reading Research Quarterly, 46,* 273–304. https://doi.org/10.1598/rrq.46.3.4

Noroozi, O., Weinberger, A., Biemans, H. J., Mulder, M., & Chizari, M. (2012). Argumentation-based computer supported collaborative learning (ABCSCL): A synthesis of 15 years of research. *Educational Research Review, 7*(2), 79-106. https://doi.org/10.1016/j.edurev.2011.11.006

Nussbaum, E. M., & Asterhan, C. S. C. (2016). The psychology of far transfer from classroom argumentation. In F. Paglieri, L. Bonelli, & S. Felletti (Eds.), *The psychology of argument: cognitive approaches to argumentation and persuasion* (pp. 407-423). London: College Publications, Studies in Logic and Argumentation series.

Osborne, J., Erduran, S., & Simon, S. (2004), Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching, 41*, 994-1020. https://doi.org/10.1002/tea.20035

Relyea, J. E., Zhang, J., Wong, S. S., Samuelson, C., & Wui, Ma. G. L. (2022). Academic vocabulary instruction and socio-scientific issue discussion in urban sixth-grade classrooms. *The Journal of Educational Research*. https://doi.org/10.1080/00220671.2021.2022584

Resnick, L. B., Asterhan C. S. C., & Clarke S. (2015). Introduction: Talk, learning, and teaching. In L.R. Resnick, C.S.C. Asterhan, & S. N. Clarke (Eds.), *Socializing intelligence through academic talk and dialogue* (pp. 1-12). Washington, DC: American Educational Research Association. https://doi.org/10.3102/978-0-935302-43-1_1

Resnick, L. B., Asterhan, C. S. C., Clarke, S., & Schantz, F. (2018). Next Generation Research in Dialogic Learning. In G. E. Hall, L. F. Quinn, & D. M. Gollnick (Eds.), *Wiley handbook of teaching and learning* (pp. 323-338). Wiley-Blackwell. https://doi.org/10.1002/9781118955901.ch13

Reznitskaya, A., Anderson, R.C., Dong, T., Li, Y., Kim, I., & Kim, S. (2008). Learning to think well: Application of argument schema theory. In C. C. Block & S. Parris (Eds.), *Comprehension instruction: Research-based best practices* (pp. 196 – 213). New York, NY: Guilford.

Reznitskaya, A., Anderson, R. C., Mcnurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S. Y. (2001). Influence of oral discussion on written argument. *Discourse Processes, 32*(2), 155-175.https://doi.org/10.1207/s15326950dp3202&3_04

Sadler, T.D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching, 41*, 513–536. https://doi.org/10.1002/tea.20009

Sadler T.D. (2011) Situating socio-scientific issues in classrooms as a means of achieving goals of science education. In Sadler T. (Ed.) *Socio-scientific issues in the classroom:*

Contemporary trends and issues in science education, *39,* 1-9. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-1159-4_1

Sadler, T.D., & Zeidler, D.L. (2005). Patterns of informal reasoning in the context of socioscientific decision making. *Journal of Research in Science Teaching, 42*, 112-138. https://doi.org/10.1002/tea.20042

Shi, Y., Matos, F., & Kuhn, D. (2019). Dialog as a bridge to argumentive writing. *Journal of Writing Research, 11*(1), 107-129.https://doi.org/10.17239/jowr-2019.11.01.04

Soter, A., Wilkinson, I. A. G., Murphy, P. K., Rudge, L., Reninger, K., & Edwards, M. (2008). What the discourse tells us: Talk and indicators of high-level comprehension. *International Journal of Educational Research, 47* (372–391). https://doi.org/ 10.1016/ j.ijer.2009.01.001

Squire, A. & Clark, S. K. (2020). Exploring how fourth-grade emerging bilinguals learn to write opinion essays, *Literacy Research and Instruction, 59* (1), 53-77. https://doi.org/10.1080/19388071.2019.1686093

The Texas Tribune. (2017-18). *Budewig Intermediate School*. Retrieved from https://schools.texastribune.org/districts/alief-isd/budewig-intermediate-school/

Turner, S. L., & Lapan, R. T. (2005). Evaluation of an intervention to increase non-traditional career interests and career-related self-efficacy among middle-school adolescents. *Journal of Vocational Behavior, 66*(3), 516-531. https://doi.org/10.1016/j.jvb.2004.02.005

Venville, G.J., & Dawson, V.M. (2010). The impact of a classroom intervention on grade 10 students' argumentation skills, informal reasoning, and conceptual understanding of science. *Journal of Research in Science Teaching, 47*, 952-977. https://doi.org/10.1002/tea.20358

Vygotsky, L. (1962). *Studies in communication. Thought and language* (E. Hanfmann & G. Vakar, Eds.). Cambridge, MA: MIT Press. https://doi.org/10.1037/11193-000

Walker, K. A., & Zeidler, D. L. (2007). Promoting discourse about socioscientific issues through scaffolded inquiry. *International Journal of Science Education, 29*(11), 1387-1410. https://doi.org/ 10.1080/09500690601068095

Wells, G. (1999) *Dialogic inquiry: Toward a sociocultural practice and theory of education*. Cambridge, UK: Cambridge University Press.

Zhang, X., Anderson, R.C., Morris, J., Miller, B., Nguyen-Jahiel, K., Lin, T., Zhang, J., …Hsu, J. Y. (2016). Improving children's competence in decision making: Contrasting effects of collaborative interaction and direct instruction. *American Educational Research Journal, 53* (1), 194-223. https://doi.org/ 10.3102/0002831215618663

Zhang, J., Niu, C., Munawar, S., & Anderson, R. C. (2016). What makes a more proficient discussion group in English language learners' classroom? Influence of teacher talk and student backgrounds. *Research in the Teaching of English, 51*(2), 183-208.

Zhang, J., Lee, G., Iluore, A. C., Relyea, J. E., & Wui, M. G. L. (2022). Fostering civic reasoning through disciplinary literacy. *The Reading Teacher*. https://doi.org/10.1002/trtr.2143

Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching, 39*, 35-62. https://doi.org/10.1002/tea.10008