Education Tech Research Dev (2018) 66:733–766
https://doi.org/10.1007/s11423-018-9574-1

DEVELOPMENT ARTICLE

Argue like a scientist with technology: the effect of within-gender versus cross-gender team argumentation on science knowledge and argumentation skills among middle-level students

Pi-Sui Hsu[1] · Margot Van Dyke[2] · Thomas J. Smith[1] · Chee-Kit Looi[3]

Published online: 23 February 2018

Abstract The purpose of this mixed-methods study was to explore the effect of within-gender and cross-gender team argumentation on seventh graders' science knowledge and argumentation skills in a computer-assisted learning environment in the United States. A total of 58 students were engaged in the collaborative within-gender team argumentation process (the treatment condition), while 46 students were engaged in the collaborative cross-gender team argumentation process (the control condition). Verbal collaborative argumentation was recorded and the students' post essays were collected. There were no statistically significant differences in science knowledge between the treatment and control conditions either for the combined set of students, or for females and males considered separately. For the combined set of male and female students, MANOVA indicated no statistically significant within-gender/cross-gender team argumentation differences in argumentation skills. Similarly, no significant within-gender/cross-gender team argumentation differences were observed among females. However, this study found a marginally significant difference in argumentation skills between male students in the within-gender team argumentation (treatment condition) and male students in the cross-gender team argumentation (control condition). A qualitative analysis was conducted to examine how the computer-assisted application supported students' development of argumentation skills

✉ Pi-Sui Hsu
phsu@niu.edu

Margot Van Dyke
mvandyke@dg58.org

Thomas J. Smith
tjsmith@niu.edu

Chee-Kit Looi
cheekit.looi@nie.edu.sg

[1] Department of Educational Technology, Research and Assessment, Northern Illinois University, 101C Gabel Hall, DeKalb, IL, USA

[2] O'Neill Middle School, 635 59th St, Downers Grove, IL, USA

[3] National Institute of Education, Nanyang Technological University, Singapore, Singapore

in within-gender and cross-gender team argumentation. Female teams, regardless of within-gender or cross-gender team argumentation, demonstrated balanced participation in the construction of argumentation maps in the application. Male teams in within-gender team argumentation (the treatment condition) demonstrated unbalanced participation in the construction of argumentation maps in the application.

Keywords Collaborative argumentation · Computer-assisted application · Middle-level students · Within-gender and cross-gender team argumentation · Graph-oriented application

Introduction

The US has invested effort and funding during the past decade to promote improvement in STEM education; nevertheless, too few American students are choosing Science, Technology, Engineering, and Mathematics (STEM) careers (Chen 2013). Furthermore, numerous studies (George 2006; Quinn and Cooc 2015; Robinson et al. 2016) have suggested that girls tend to start losing interest in science at the middle school level and the difference between boys' and girls' interests in science appear to widen as they grow older. As research has shown a positive relationship between interest in science and choice of science as a career, and classroom experience is one of the factors that influence girls' interests toward science (Hoffmann 2002; Kelly 1978; Welch et al. 2014), this study aimed at promoting girls' interests in science by enriching girls' classroom experience through involving them in collaborative science discourse.

Recently, the *Next Generation Science Standards* (NGSS), built on *A Science Framework for K-12 Science Education* (National Research Council 2012), identified "engaging in argument from evidence" (p. 12) as one of the essential eight science practices for students. Scientific argumentation is a form of logical discourse of arriving at an agreed-upon position among members of a group (Andriessen 2006; Driver et al. 2000) and is practiced when scientists build on and refute one another's theories and empirical evidence to arrive at scientific conclusions, while engaging in a debate process where the intention is to prevail over an opponent (Kuhn 1993; Walton 1996). While Sampson et al.'s (2012) view of scientific argumentation is consistent with earlier views (e.g., Andriessen et al. 2003), they expanded the definition of scientific argumentation and view it as a social and collaborative process of proposing, supporting, evaluating, and refining ideas in an effort to make sense of a complex problem to advance knowledge. Collaborative scientific argumentation in this study is a critical element to support the development of individual argumentation skills (Andriessen 2006; Cho and Jonassen 2003; Crowell and Kuhn 2014; Hogan et al. 1999; Jonassen and Kim 2010; Kim et al. 2007; Kuhn and Udell 2003). It also supports students in developing knowledge and a deeper understanding of how scientific knowledge is generated (Manz 2015).

A number of gender studies (Carlone et al. 2014; Halpern et al. 2007) have pointed out that, at puberty, girls move from self-confidence to self-consciousness, start to censor themselves, and become increasingly silent in class. However, these studies have also shown that girls who work collaboratively are less likely to suffer loss of interest in science. This present study viewed computer-supported collaborative learning environments as a potential way to engage girls in scientific argumentation and aimed at orchestrating a computer-supported collaborative learning environment by exploring the use of a graph-oriented computer-assisted application and gender grouping strategies to scaffold students' collaborative scientific argumentation.

In the following section, we will discuss potential benefits of the use a graph-oriented computer-assisted application to support argumentation and content knowledge. We also will discuss the rationale of the gender grouping strategies intended to encourage the involvement of girls in the collaborative scientific argumentation process.

Use of a graph-oriented computer-assisted application to support argumentation and content knowledge

In light of the reform efforts, researchers have used different approaches to develop curricula to help middle-level students develop scientific argumentation skills (Evagorou and Osborne 2013; Iordanou 2010; Kuhn et al. 2008, 2010; Kuhn 2011, 2015). For example, Kuhn and Udell (2003) investigated whether peer dialogue was effective in developing argumentation skills among 34 13- to 14-year-olds who were academically at-risk. The students in the peer dialogue group showed increased frequency of use of powerful argumentation skills and improved quality of individual arguments when compared to the control group. Crowell and Kuhn (2014) developed a curriculum in which 56 students (sixth, seventh, and eighth grades) in an urban middle school participated twice a week for 3 years. The Internet chatting application, Google Chat, supported the argumentation curriculum in the experimental group, while the control group participated in a traditional whole-class discussion. Argumentation skills of the experimental group outdistanced those of the control group.

More recent studies (Dwyer et al. 2012a, b; Hsu et al. 2015, 2016; Scheuer et al. 2010) have explored the potential of graph-oriented computer-assisted applications (GOCAA) and found a positive impact on argumentation skills. Research shows that visualizing arguments graphically through a graph-oriented computer-assisted application enables students to see the structure of the argument, thus facilitating more rigorous construction and communication (Kiili 2012; Scheuer et al. 2010). There are several graph-oriented applications (e.g., Digalo, Belvedere, Araucaria), each of which typically has a distinct way of constructing argumentation maps. However, there are many features common across these applications. For example, contributions are displayed as boxes or nodes that represent argument components. Texts can be inserted into the boxes or nodes to represent argumentation. Arrows represent relationships among the argument components (e.g., supports or refutes). As the different components of arguments and their relationships can be distinguished via their visual appearance, learners are able to visualize and identify the important ideas in argumentations as concrete objects. These objects can then be pointed to, linked to other objects, and discussed. Figure 1 shows the GOCAA used in this study.

A GOCAA is an example of a cognitive tool (Pea 1985). Such an application has potential for assisting learners in accomplishing cognitive tasks and leads to the development of argumentation skills. Cognitive tools can serve four purposes: (a) supporting cognitive processes such as memory and metacognitive processes; (b) sharing cognitive load by providing support for lower level cognitive skills so that resources are left for higher order thinking skills; (c) supporting engagement in cognitive activities such as global databases, that otherwise would be out of reach; and (d) generating and testing hypotheses (Lajoie and Derry 2013). These four purposes are not mutually exclusive. In a number of studies (Carr 2003; Dwyer et al. 2012a, b; Easterday et al. 2009), a GOCAA served the first and second purposes to support the development of argumentation skills. Dwyer et al. (2012a, b) conducted a study in which 74 undergraduate psychology students

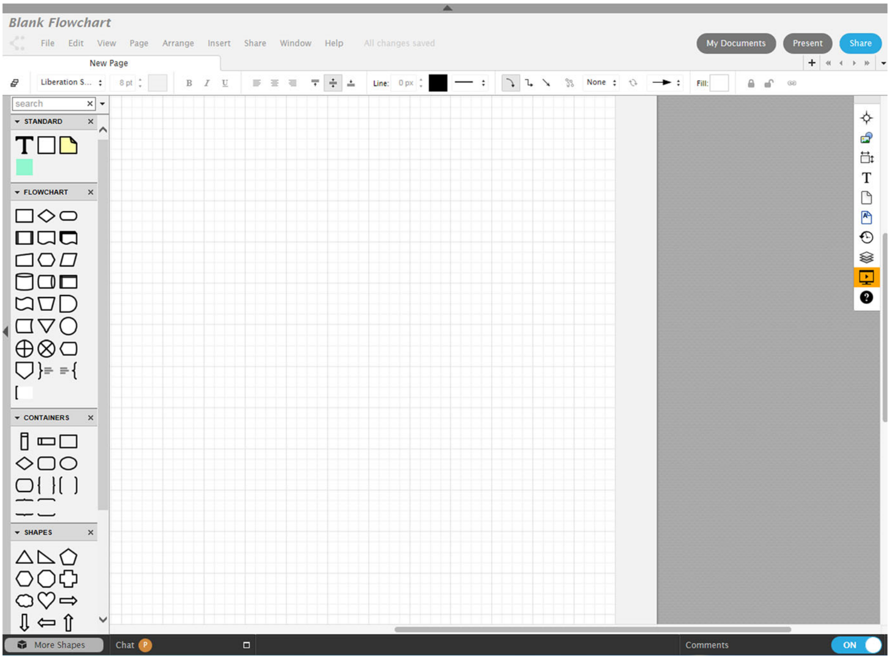


Fig. 1 Screenshot of Lucidchart

were allocated to conditions in which the infusion of a graph-oriented computer-assisted e-learning environment was either present or absent. The students in the infused environment showed significantly larger gains in argumentation analysis than the control group. In their study, the students used the GOCAA over 8 weeks, both inside and outside of the classroom, to build their own argumentation. The visual representation of a GOCAA supported cognitive processes and allowed the students to make their thinking visible (memory) and monitor the development of argumentation (metacognition). Additionally, the visual representation also facilitated the sharing of cognitive load by providing support for lower level of argumentation and provided the students with more resources for higher level of argumentation.

With the general positive impact of a GOCAA on argumentation skills, several studies (Carr 2003; Easterday et al. 2009) have explored the potential of a GOCAA to develop content knowledge, with mixed findings. Easterday et al. (2009) conducted a study in which GOCAAs were used to teach causal reasoning on public policy problems. The study compared the effects of three conditions under which students were asked to analyze a problem. These conditions included: (a) problem presented as text only; (b) problem presented as text with an additional pre-made causal diagrams; (c) problem presented as text with a GOCAA that students could use to actively construct a diagram from the text. Scores on the transfer test were significantly better for students in the third condition. Carr (2003), however, had contrary findings indicating that a graph-oriented computer-supported environment is not necessarily better than traditional methods for promoting learning outcomes. In Carr's study, second-year law students in a treatment group worked in small study groups of on legal problems while having access to a GOCAA, QuestMap, while students in a control group worked without QuestMap, either alone or in small groups. The students in the treatment group did not outperform those in the control group on a final exam.

With mixed findings from earlier studies (Carr 2003; Dwyer et al. 2012a, b; Easterday et al. 2009; Schwarz and Glassner 2007; Suthers and Hundhausen 2003; Suthers et al. 2008), we identified a number of key elements that require both clear descriptions and further research, which might facilitate a better understanding of the mixed findings pertaining to the impact of a GOCAA to support the development of content knowledge. The first area involves the clear orchestration of the sequence of activites when using a GOCAA to mediate the collaborative argumentation construction process. One set of studies (e.g., Carr 2003; Weinberger et al. 2010) used a GOCAA to support synchronous collaborative argumentation in class. In Carr's study, for example, the second-year law students began constructing an argument, saw another counterargument appear, and then began discussing the counterargument with its creator outside the GOCAA, QuestMap in a treatment group. This study did not provide a description about the sequence of activities involved in the collaborative argumentation process, nor the amount of time allocated for each activity. It also is unclear whether there were any prompts in the GOCAA to scaffold the process. Without careful planning of the activities to facilitate the process, the students might have used the tool for transcription assistance instead of collaborating with one another through the application. Another set of studies (e.g., Strijbos 2011; Suthers et al. 2008) used a GOCAA to support solely asynchronous collaborative argumentation. In Suthers et al.'s quasi-experimental study, for example, pairs of college- level physics majors from various geographic locations synchronously collaborated on a science challenge problem using an online application. Although the authors provided clear description of the sequence of the activities, the study was conducted in a controlled-experimental setting and the time span was short (e.g., 120 min). The present study intends to address these issues and examine how the activities could be orchestrated into daily classroom experience, and also intends to explore the impact of the graph-oriented computer-assisted application on students' development of science knowledge and argumentation skills in the middle-level science classroom.

The second issue concerns how groups are formed. Some studies (e.g., Carr 2003) have used small groups of students for graph-oriented, computer-assisted, collaborative argumentation activities. In these studies, however, information pertaining to the number of students in each group is unclear, as information about the number of female and male students in each group. More recent studies (e.g., Suthers et al. 2008) have addressed some of the issues in forming groups for collaborative argumentation activities, such as gender distribution. However, some grouping strategies are more practical for college-level students than middle-level students. For example, Suthers et al. used pairs of college-level students in the asynchronous collaborative argumentation process. To avoid social awkwardness, they assigned pairs of students based on acquaintances, regardless of gender composition, to work on the same problems. They did not report any issues pertaining to college-level students working in pairs. On the contrary, however, Ding et al. (2009, 2011) showed that having young adolescents work in mixed-gender pairs can lead to unproductive communication. As such, in the present study, we surmised that forming pairs of students might not be a feasible grouping strategy for young adolescents in a computer-assisted collaborative learning environment. Therefore, rather than using pairs of students, the present study employed groups of 3–4 middle-level students to engage in the collaborative argumentation process. The following section explicates the rationale for the different types of gender groupings used in this study.

Impact of types of gender groupings on male and female students' learning in computer-assisted collaborative learning environments

In the computer-assisted collaborative learning literature (Sullivan et al. 2015; Zhan et al. 2015), gender differences in communication styles and cognitive abilities play an important role in the influence of gender groupings on male and female students' learning outcomes. For example, Ding et al. (2009, 2011) studied whether gender differences were evident in the effect of gender grouping (mixed- vs. single-gender pairings) in a computer-assisted collaborative learning environment in a secondary school. Pairs of students participated in a collaborative activity over a period of 2 weeks to solve physics problems. In each pair, students use an Internet-based computer program to communicate with their partner. The program provided a computer-assisted collaborative learning environment through which each student could use text messages and pictorial messages in the text-messaging box to communicate. Analysis of 96 secondary students' interactions revealed that a divergent pattern of knowledge elaboration led to female students' showing higher learning outcomes in single-gender pairs than in mixed-gender pairs. This observed effect of single-gender pairing, however, was not evident for males. Specifically, the authors observed that, when a female student in a pair was still stuck on the force analysis, her male student had already started calculation. Additionally when a female complained that her male partner moved too fast, the male partner ignored her complaints and continued with his calculation. The female then abandoned her questioning and accepted her male partner's answer. Such behavior could lead to a divergent pattern of knowledge elaboration for female students. Research on gender differences in communication styles has shown that, while females use communication as a tool to enhance social connections and create relationships, males communicate to exert dominance and achieve tangible outcomes (Leaper 1991; Maltz and Borker 1982; Wood 1996). With communicate styles aiming to achieve tangible outcomes, male students tend to lose sight of the importance of bringing their female partner to the same tempo to overcome the lack of shared context in the computer-supported learning environments, which might cause a divergent pattern of knowledge elaboration for females.

Similarly, a number of studies (Brotman and Moore 2008; Prinsen et al. 2007; Zhan et al. 2015) showed females in single-gender groups tend to perform better than those in mixed-gender groups in a computer-supported collaborative learning environment. In Zhan et al.'s (2015) study, a sample of 588 undergraduate students enrolled in a digital design course were randomly assigned to 147 four-student groups that fell into five categories according to the composition of group members' gender. The five categories were designated as 4 M (four males), 3M1F (three males and one female), 2M2F (two males and two females), 1M3F (one male and three females), and 4F (four females). Results indicated that the 2M2F and 4F groups significantly outperformed the other groups. This may have been due to the fact that female students show better planning and communication skills than males (Korobov 2013; Stokoe 2004; Tarim and Kyratzis 2012). Additionally, female students may engage in more discussion in computer-supported collaborative learning than males, ultimately leading to a better outcome. Several studies (Asterhan et al. 2012; Caspi et al. 2008; Ma and Yuen 2011) have reported that female students demonstrate higher participation than males in computer-assisted collaborative activities. One of these studies (Asterhan et al. 2012) examined single-gender groups. In this study, 82 ninth graders discussed social issues in a graph-oriented computer-assisted program. Compared to homogeneous all-boy groups, homogeneous all-girl groups had higher degree of participation (textual contribution) and scored higher on aspects of collaborative argumentative quality. To conclude, more interaction appears to occur in female-only groups than in male-only groups.

Although these studies found that homogeneous female-only groups primarily benefit females, these studies pointed out that males in mixed-gender and gender-balanced groups tend to perform better than males in homogeneous single-gender groups. Male students may benefit more than females from the dynamic atmosphere created by mixed-gender communication, leading to better cooperation and enabling them to achieve better learning outcomes than male students in single-gender groups. Additionally, male students may perform better than females in a gender-balanced environment due to a lower level of power issues, where no individual student feels he is in a minority. This also reduces the likelihood of freeloading.

In addition to different communication styles, a number of studies (Baram-Tsabari and Yarden 2011; Ding et al. 2009, 2011) have shown that female and male students have differences in cognitive abilities that may impact how they elaborate knowledge in a computer-assisted collaborative learning environment. For example, in Ding et al. (2011), males in mixed-gender pairs tended to use visual representation to answer their female partners' questions. By contrast, female students tended to use text-based messages and verbal explanations. Buck et al. (2012) also found that female students are more willing to express their ideas while males tend to ask for information. Female students are more verbal than males and tend to share whatever they think might be relevant with their group members, while males tend to share only task-related information. Gender differences in cognitive abilities have been widely analyzed in the psychological and neuropsychological literature (e.g., Hyde and Linn 1988; Buffery and Gray 1972; Caplan et al. 1997; Halpern 1992; Fairweather 1976) and major differences in cognitive abilities between males and females typically have been reported in: (1) verbal abilities, favoring females and (2) spatial (visual) abilities, favoring males.

In light of gender differences in communication styles and cognitive abilities described above, this study surmised that when coupling verbal argumentation with female team members and engaging in graph-oriented computer-assisted application that allowed female students to use their textual (verbal) communication with another female team, female students would benefit most from within-gender team argumentation in the collaborative argumentation process. However, with the advance of technology, both female and male students have more opportunities to engage in collaborative learning activities (Abnett et al. 2001; Chung et al. 2013). It is reasonable to expect that female and male students need to engage in some form of science talk in the classroom. Thus, we also were interested in exploring whether female students would perform worse when engaging in graph-oriented computer-assisted application with the male team (cross-gender team argumentation) than those in the within-gender team argumentation. Additionally, research has explored female and male students' learning and interests in science and has attempted to make curriculum and pedagogy more inclusive for both genders (Velayutham et al. 2012). We also explored how male students performed in both conditions. In previous gender grouping studies, when research refers to mixed-gender groupings, it typically suggests that female and male students are mixed to form a team. In this study, because all teams were formed with students of the same gender, we use the term, "cross-gender team argumentation," to refer to a homogeneous-gender team arguing with another homogeneous-gender team of the opposite gender (e.g., female team vs. male team).

The purpose of this mixed-methods study was to explore the effect of within-gender and cross-gender team argumentation on seventh graders' science knowledge and argumentation skills in a computer-assisted learning environment in the United States. This project addressed critical STEM educational needs by studying the essential practice of scientific argumentation, and contributed to the literature in the computer-supported collaborative learning environment.

Research questions

The following research questions were addressed:

1. What are the differences in science knowledge (as measured by scientific facts, scientific explanations, and valid scientific facts/explanations) between students in within-gender team argumentation (treatment condition) and students in cross-gender team argumentation (control condition) in a GOCAA learning environment?

    1a. What are the differences in science knowledge between female (or male) students in within-gender team argumentation and female (or male) students in cross-gender team argumentation in a GOCAA learning environment?

2. What are the differences in argumentation skills between students in within-gender team argumentation and students in cross-gender team argumentation in a GOCAA learning environment?

    2a. What are the differences in argumentation skills between female (or male) students in within-gender team argumentation and female (or male) students in cross-gender team argumentation in a GOCAA learning environment?

3. What are the differences in argumentation maps in Lucidchart between female teams (or male teams) in within-gender team argumentation and female teams (or male teams) in cross-gender team argumentation?
4. If there was a difference in argumentation skills, in what ways would the graph-oriented computer-assisted application support female and male teams' development of argumentation skills in different types of gender groupings?

In research questions 1 and 2, we were interested in exploring the effect of gender groupings on all students between two conditions. In research question 1a and 2a, we were particularly interested in exploring the effect of gender groupings on female and male students between two conditions.

Methods

Research design and participants

This mixed-methods study (Creswell 2013) was conducted in a seventh grade middle school science classroom in suburban Chicago, U.S. Mixed-methods research design involves the process of collecting, analyzing, and mixing both quantitative and qualitative data in a single study for a better understanding of a research problem. In this study, the quantitative aspect consisted of analysis of scores pertaining to science knowledge and argumentation skills obtained from post essays, and analysis of the frequency of each argumentation skill on argumentation maps in a GOCAA (Lucidchart), whereas the qualitative aspect involved exploring the role of the GOCAA in the collaborative argumentation process. Therefore, by using a mixed-methods approach, this study allowed for increased depth of explanation when answering the research questions. There were six classes participating in the study. A total of 58 students (29 females and 29 males, 3 classes) comprised the treatment (within-gender team argumentation) condition while a total of 46 students (24 females and 22 males, 3 classes) were in the control (cross-gender team argumentation) condition. The composition of the students' ethnic background was diverse. Half (50%) came from middle-class Caucasian families, 25% were from middle-class Asian American families and the remaining 25% were from middle-class African-American or Hispanic-American families. The diversity of ethnic background was approximately uniformly distributed across classes. The students' science performances ranged from low to high. The same science teacher taught all students. The teacher had taught science for more than 10 years in the middle school, and was instrumental in integrating technology into instruction. Thus, the students had already learned about different types of technology and had participated in various technology-supported or computer-supported activities.

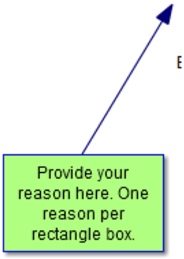
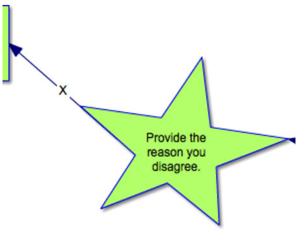
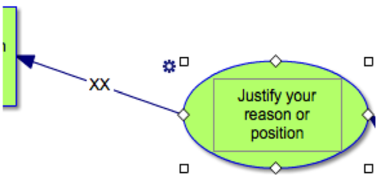
Each of the six classes was randomly assigned to either the treatment (within-gender team argumentation) or control condition (cross-gender team argumentation). In both conditions, the students worked in teams of three to four. In within-gender team argumentation (treatment condition), each team was engaged in verbal collaborative argumentation with the members of their homogeneous-gender (e.g., all-female) team and then argued with another other team whose gender composition was the same as theirs (e.g., all female students) using the graph-oriented computer-assisted application (described in the next section). There were 11 female teams. Typically, one female team was assigned to argue with the other female team. However, there was an odd number (11) of teams. Among 11 female teams, three female teams were assigned to argue with each other. Likewise, one male team was assigned to argue with the other male team. There was an odd number (9) of male teams. Among 9 male teams, three male teams were assigned to argue with each other. There were 6–7 teams per class in the treatment condition.

In cross-gender team argumentation (control condition), the teams also engaged in verbal collaborative argumentation with members of their homogeneous-gender (e.g., all-female) team, and then argued with a different-homogeneous-gender team (e.g., all-male team) using the graph-oriented computer-assisted application. In the control condition, there were 8 female teams and 8 male teams. There were 4–6 teams per class.

Graph-oriented computer-assisted application

In both treatment and control conditions, each team used the graph-oriented computer-assisted application, Lucidchart, to present their arguments and argue against a team (Table 1). We reviewed the literature (Scheuer et al. 2008, 2010) and investigated an appropriate GOCAA for this study. We used a number of criteria to select a potential application, such as capacity to support argumentation and expressiveness, allowance for online synchronous/asynchronous collaboration, and compatibility with PC and Mac platforms. Most applications were outdated (e.g., CoFFEE) or were designed for college students (e.g., Belvedere). We selected and tailored Lucidchart to meet the needs of this study. Lucidchart is a propriety tool and is developed for various learning purposes. A handout that indicates argumentation elements by corresponding shapes and arrows and definition (Kuhn 1993) was provided for each student in this study.

Table 1 Argument elements by corresponding shapes and arrows in Lucidchart and definitions

Shapes and arrows	Argumentation skill	Definition
	Position (light bulb)	An opinion or conclusion on the main question
	Reason (rectangle and arrow)	A claim that supports the position
	Evidence (cloud and arrow)	A separate idea or example that supports a reason (or counterclaim or rebuttal)
	Counterargument (signified by star and "x")	A claim that refutes another position or gives an opposing reason
	Rebuttal (signified by oval and "xx")	A claim that refutes a counterargument by demonstrating that it is invalid, lacks as much force or correctness as the original argument or rests on a false assumption

Instruments

We developed a rubric to assess science knowledge and a scoring method to assess students' scientific argumentation skills on their individual essays. A sample student post essay is presented in Appendix 1.

Science knowledge rubric

To score students' science knowledge post essays on alternative energy, a rubric based on the NGSS (National Research Council 2012) was developed (Appendix 2). Specifically, the rubric focuses on one of the Earth and Spaces sciences standards (Earth and Human Activity: MS-ESS 3.3) at the middle school level, examining how students assess the kinds of solutions that are feasible and can reduce human impact on the environment. The rubric was validated by two middle school science teachers and two college level science educators. The three measures of science knowledge were based on (a) the number of scientific facts presented, (b) the number of scientific explanations provided, and (c) the scientific validity of the facts and explanation.

Based on the post essays, each student received a score on each measure that could range from 0 to 3. The scores of three measures were then totaled and recorded. As one example of how a student essay was scored, a student pointed out the fact that "Wind energy is a clean renewable resource" and used the explanation "Wind turbines don't have to burn anything to generate electricity" to support this fact. In this case, this student earned 1 point for presenting a scientific fact and 1 point for providing a statement of scientific explanation to support it. Scientific validity was deemed present when the facts presented in the essays were, in fact, correct. When all the facts were correct, 3 points were earned. If there was one false fact, 2 points were earned. Student examples of false facts/explanation included "Biomass mostly comes from trees and that's something the U.S. won't run out of, for sure," and "Only a couple of birds die every year from wind turbines."

Argumentation skill scoring method

Based on Kuhn's (1993) definition of individual argumentation skills, the students' essays were scored for argumentation skills (reason, evidence, counterargument and rebuttal skills). A *position* refers to an opinion or conclusion on the main question that is supported by *reason*. *Evidence* is a separate idea or example that supports reason or counterargument/rebuttal. *Counterargument* refers to an assertion that counters another position or gives an opposing reason. A *rebuttal* is an assertion that refutes the counterargument by demonstrating that the counterargument is not valid, lacks as much force or correctness of the original argument, or is based on a false assumption. The students were required to follow correct logic to receive scores for each argumentation skill (Appendix 3). For example, a reason must follow a position. An evidence must follow a reason. A counterargument must follow a reason and evidence. A rebuttal must follow a counterargument. When the students presented one reason (e.g., "solar energy can be used anyplace.") on their essay, they received 1 point. The same scoring procedure applied to evidence (e.g., "solar panels can be installed everywhere"), counterargument (e.g., "If it is cloudy, it probably won't work"), and rebuttal skill (e.g., "My mom told me one time when I was swimming that it is easier to get burned when it is cloudy because the sun rays peek through the clouds when you don't even know it"). Each student's essay was scored individually. Each student received four argumentation scores and the total score for each argumentation skill was recorded in SPSS. The minimum score for each argumentation skill was 0. The researchers did not limit the maximum score for each argumentation skill.

Procedure/data collection

Beginning of the semester

The science teacher surveyed all students about their prior knowledge about alternative energy and prior argumentation experience. A few students knew the term "alternative energy," but all students had no prior topic knowledge about alternative energy. Similarly, all students had no prior argumentation experience. The science teacher surveyed all students about their prior knowledge in the use of the GOCAA or similar tools. A few students knew the term "concept-mapping tool," but they had never used the GOCAA or any similar tool. The topic was new to the seventh graders and was part of the middle school curriculum. Therefore, at the beginning of the school year, the students in both conditions learned both collaborative argumentation skills and the use of the GOCAA, Lucidchart. In the middle of the fall semester, all students in both conditions researched the assigned topic of alternative energy for 2 weeks and developed an iMovie video clip to present their findings in a team of three to four persons. The potential sources of energy included solar, biomass, geothermal, hydrogen, hydropower, wind, and nuclear. The project was developed based on the U.S. NEED (National Energy Education Development) initiative Each team researched one form of alternative energy. They were able to learn about other forms of alternative energy from other teams' presentations. All students were aware that they needed this knowledge to participate in the collaborative argumentation activity. They took notes while listening to other teams' presentations. The science teacher provided detailed instructions regarding what content they needed to include in the presentations. The science teacher verified the correctness of the students' content before they made the presentations. Students did not necessarily work in the same teams that would later develop collaborative argumentation.

The collaborative argumentation activity lasted for 1 week. On each day, the first two researchers provided brief instructions for the activity in both conditions and addressed students' questions. The researchers designated a leader in each team and the leader was responsible for facilitating the logistics in the process, such as picking up and returning the computer. The researchers explained the difference between dominating and facilitating and ensured that every student could contribute. In both conditions, the first two researchers constantly rotated among the teams and monitored their progress. Table 2 outlines the daily argumentation activities in both conditions and the following paragraphs explain the activities in more detail.

Collaborative argumentation activity day 1 and 2

During the first 2 days, the first two researchers instructed students to argue their position, reason, and evidence with their team members. All students participated in the collaborative argumentation activity. Students in both conditions were given 40 min each day in class to engage in intra-team verbal collaborative argumentation with the members of their homogeneous-gender team (e.g., all female students or all male students) and argued "If the US could fund only one form of alternative energy, which one should you select?"

Table 2 Daily argumentation activities in the treatment and control conditions

	Treatment condition (within-gender team argumentation)	Control condition (cross-gender team argumentation)
Day 1 and 2	Each team member started engaging in verbal argumentation with their team members Each team used the GOCAA (Lucidchart) to post position, reasons, and evidence	
Day 3, 4, and 5	Each female team argued with the other female team in the GOCAA (Lucidchart) Each male team argued with the other male team in the GOCAA (Lucidchart).	Each female team argued with the other male team in the GOCAA (Lucidchart)
Day 6	Each student completed a post-essay	

After each team came to a consensus about which form of alternative energy should be funded by the U.S. (position), each team used the graph-oriented computer-assisted application, Lucidchart, to post their position, reasons and evidence (Fig. 2).

Collaborative argumentation activity day 3, 4 and 5

Starting on the third day, inter-team argumentation was initiated in Lucidchart. The inter-team argumentation took place for a 40-min period for 3 days. Each team was paired with a corresponding team in the same class. Each team was aware of the corresponding team's members and their gender because the science teacher shared the team assignment with the class before the argumentation activities started. The first two researchers instructed students to provide counterarguments toward the other team and provide rebuttals to defend their position. In the treatment condition (within-gender argumentation), each team read the reasons and evidence of a team whose gender composition was identical to its own (e.g., a male team read reasons and evidence of another male team), and provided counterarguments in Lucidchart. Then each team read the counterarguments, continued to argue how to rebut the counterarguments, and posted their rebuttals in Lucidchart. Figure 2 indicates a sample argumentation map constructed by two male teams in the treatment condition. Each team indicated its position by inserting its position below a light bulb. During class time each team used the shapes and arrows on Table 1 to represent its argumentation and argue with other team through Lucidchart. Team four selected nuclear (orange) and team five selected hydropower (blue). As indicated in Fig. 2, for journal printing purposes, we converted the shapes in orange to plain textboxes with a thicker border, and the shapes in blue to plain textboxes with less thick border. Each team used the shapes and arrows (shown on Table 1) to represent its argumentation and argued with the opposing team in Lucidchart. In the control condition, each team argued with the other team whose gender composition differed from its own (e.g., an all-female team arguing with an all-male team) in Lucidchart, using counterargument and rebuttal. In both conditions, each team was designated a quiet corner in the spacious library or in the classrooms.

This study attempted to have minimal researcher intervention. Thus, the first two researchers only intervened when students were off-task or off-topic in both conditions. Although the researchers provided brief instructions at the beginning of each class, it was not necessary for the students to have followed these instructions. The purpose of this study was to understand the dynamics of collaborative argumentation process with the support of

Fig. 2 Sample argumentation map in Lucidchart constructed by two male teams in the treatment condition

the graph-oriented computer-assisted application. As long as they stayed on track, we did not force students to follow the instructions. For example, some teams started to argue about counterarguments during the first 2 days and the researchers did not force them to focus only on the position, reason, and evidence.

Collaborative argumentation activity day 6

After a week (5 school days), the students in both conditions were asked to write post essays. The topic was, "Which form of alternative energy is the best?" The students did not presume that "the best" meant the same as "Which form of alternative energy would you fund if you could only fund one?" They were give 40 min to write their individual essay. There were no page limitations for their essay. They could refer to the argumentation maps in Lucidchart and any materials including their own notes. The science teacher collected their essays and did not provide any opportunity for them to revise.

Quantitative analysis of essays for science knowledge and argumentation skills

There were a total of 104 student post essays. The students' scores for individual essays were not counted into the class grades. There were two raters for science knowledge, the first and second researcher. Both raters had expertise in the NGSS and in alternative energy. Before both raters began to score all essays, they randomly selected five essays and scored each essay independently. The interrater reliability (percentage agreement) for these five essays was 80%. The raters shared their ratings and discussed discrepancies. They then continued to score the essays. The interrater reliability (percentage agreement) for the full set of essays was 88%.

As with the ratings of scientific fact, explanation, and validity, there were two raters of argumentation skill, the first researcher and a doctoral student. Both raters had expertise in argumentation skills. Before the raters began to score the complete set of essays, they randomly selected five essays and scored each independently. The interrater reliability (percentage agreement) was 85%. The raters then shared their ratings and discussed discrepancies. After resolving the discrepancies, the raters continued to score all essays. The interrater reliability (percentage agreement) for the complete set of essays was 90%.

Quantitative analysis of collaborative argumentation maps in GOCAA for argumentation skills

The researchers conducted analysis of each team's collaborative argumentation map in Lucidchart for argumentation skills. There were 9 argumentation maps in Lucidchart in the treatment condition. There were eight argumentation maps in the control condition.

The researchers analyzed each argumentation map in Lucidchart and identified whether each team used the correct shape and arrow to represent each argumentation skill. The researchers assigned the teams that made only one (e.g., incorrect color, incorrect shape) or none errors as good alignment between argumentation skills and corresponding shapes/arrows. The researchers also tallied the number of shapes that represent each argumentation skill and, for each of the argumentation skills, compared this number for female and male teams between the treatment and control conditions. The interrater reliability (percentage agreement) was 95%.

Qualitative analysis of verbal collaborative argumentation process for the role of the graph-oriented computer-assisted application

Kelly and Crawford (1996) developed the framework to analyze how the computer-assisted application supports the interaction among learners. In this study, we modified it to analyze how the GOCAA supports the collaborative argumentation process. See Table 3 for the coding scheme.

To help explain our coding process, the researchers selected an excerpt that involved students' use of Lucidchart to support their verbal collaborative argumentation in a treatment team. Table 4 shows examples from three speakers (Alan, Ben, and Celine) arguing using Lucidchart (see Fig. 2), along with the GOCAA's action, the nonverbal actions of three speakers, and the initial, researcher-assigned codes for the development of argumentation. The unit of analysis was idea units. Bransford and Johnson (1973) have defined an idea unit as "corresponding either to individual sentences, basic semantic propositions, or phrases" (p. 393). The researchers began looking for idea units by

Table 3 Coding schemes

Categories	Code	Definition	Examples
Functions of computer application			
Constructing	Cons	Coding as a graph-oriented computer-assisted application is used to show learners' position, reasons or evidence	A graph-oriented computer-assisted application is used to represent a position, reasons or evidence by inserting different shapes
Exhibiting	Exh	Coding as a graph-oriented computer-assisted application is used as external representation	A graph-oriented computer-assisted application shows different shapes to represent different argumentation skills
Eliciting	Eli	Coding as a graph-oriented computer-assisted application serves as external representation to stimulate more responses from learners	A graph-oriented computer-assisted application shows position to learners and learners respond by providing more reasons and evidence
Acting as ally	Act	Coding as a graph-oriented computer-assisted application is used by learners to support their efforts to make a case	A graph-oriented computer-assisted application is used as learners' ally to provide counterarguments or rebuttals
Types of affordance in the argumentation process			
Demonstrating	Dem	Coding as learners use different shapes to represent different argumentation skills	Learners post position, reasons, and evidence on a graph-oriented computer-assisted application
Reading	Rea	Coding as learners use the external representation of a graph-oriented computer-assisted application to make sense of the process	Learners read reasons and evidence provided by their corresponding team
Responding	Res	Coding as learners provide more reasons and evidence to support their position	Learners post more reasons and evidence to support their position on a graph-oriented computer-assisted application
Claiming	Cla	Coding as learners provide counterarguments or rebuttal to make their case	Learners post counterarguments or rebuttals on a graph-oriented computer-assisted application
Regulating metacognition	Regume	Coding as learners consistently monitor and regulate the learning process by looking at the external representation	Learners reflect on the argumentation process by looking at the argumentations on a graph-oriented computer-assisted application

examining sentences in the transcriptions. Please refer to the first column in Table 4 for the sentences that we analyzed. When we assigned the initial codes, we focused on how argumentation skills were developed. This study examined the team members' interaction in the collaborative argumentation process. The researchers transcribed a total of 36 teams' video clips of verbal argumentation process supported by Lucidchart and coded them. Kuhn (2015) addressed the lack of the in-depth analysis of the group interaction as a common issue in the collaboration studies. The method used in this study was to address this issue. The interrater reliability (percentage agreement) for the first five videos was 80%. The raters then shared their ratings and discussed discrepancies. After resolving the discrepancies, the raters continued to analyze all video clips. The interrater reliability (percentage agreement) for the complete set of videos was 90%.

Table 4 A transcript of three male students working together with Lucidchart and corresponding computer's action, nonverbal action of three speakers, and the researcher-assigned codes

Students' argumentation	Computer's action	Nonverbal action	Codes
Alan (M): Nuclear is cool. Our position is nuclear. I need a reason	Allows students to use different shapes to represent a reason and evidence	Inserting a position in Lucidchart	Demonstrating a position (dem)
Alan (M): It can be built anywhere	Allows students to use different shapes to represent a reason	Inserting a reason in Lucidchart	Demonstrating a reason (dem)
Ben (M): Yes	Allows students to use different shapes to represent a reason	Inserting a reason in Lucidchart	Demonstrating a reason (dem)
Celine (M): Let's look for the stuff there (Lucidchart)	Allows students to use different shapes to represent a reason	Inserting a reason in Lucidchart	Demonstrating a reason (dem)
Alan (M): Let's see what they (hydropower team) are putting	Shows the reason provided by the hydropower team	Commenting on the screen	Reading other team's position, and reason (red)
Alan (M): Do you guys have anything (reason) else?	Shows the reason	Looking at a position and reason in Lucidchart	Generating a reason (res)
Ben (M): They are other reasons	Shows the reason	Looking at a position and reason in Lucidchart	Generating a reason (res)
Alan (M): There are several (nuclear power plants) in Illinois	Allows students to use different shapes to represent an evidence	Inserting an evidence in Lucidchart	Demonstrating an evidence (dem)
Celine (M): 6?	Allows students to use different shapes to represent an evidence	Inserting an evidence in Lucidchart	Demonstrating an evidence (dem)
Alan (M): Let's get information from iPad	Shows the reason and evidence	Looking at a position, reason, and evidence in Lucidchart	Generating a reason and evidence (res)
Alan (M): Unlimited. It creates jobs	Allows students to use different shapes to represent a reason and evidence	Inserting a reason, and evidence in Lucidchart	Demonstrating a reason and evidence (dem)
Ben and Celine: Yes	Allows students to use different shapes to represent a reason	Inserting a reason in Lucidchart	Demonstrating a reason (dem)

Table 4 continued

Students' argumentation	Computer's action	Nonverbal action	Codes
Alan (M): Low emission	Allows students to use different shapes to represent a reason	Inserting a reason in Lucidchart	Demonstrating a reason (dem)
Ben (M): Maybe require less raw materials	Allows students to use different shapes to represent a reason	Inserting a reason in Lucidchart	Demonstrating a reason (dem)
Celine (M): It gets easy transportation	Shows the reason and evidence	Looking at a position, reason, and evidence in Lucidchart	Generating a reason and evidence (res)
Alan and Ben (M): What do you mean?	Shows the reason and evidence	Looking at a position, reason, and evidence in Lucidchart	Generating a reason and evidence (res)
Alan (M): It's weather reliant	Allows students to use different shapes to represent a reason	Inserting a reason in Lucidchart	Demonstrating a reason (dem)
Alan (M): We need evidence to support this (Pointing to a reason in Lucidchart)	Shows the reason and evidence	Looking at a position, reason, and evidence in Lucidchart	Generating an evidence (res)
Alan, Ben and Celine review four reasons in Lucidchart and brainstorming	Shows the reason and evidence	Looking at a position, reason, and evidence in Lucidchart	Summarizing and reflecting (regume)
Alan (M) looking for the cloud bubble What is the ratio?	Allows students to use different shapes to represent an evidence	Inserting an evidence in Lucidchart	Demonstrating an evidence (des)
Ben (M): 28 grams of uranium releases as much as energy as 100 tons of coal	Allows students to use different shapes to represent an evidence	Inserting an evidence in Lucidchart	Demonstrating an evidence (des)
Alan (M): So that supports "Requires little raw material"	Allows students to use different shapes to represent a reason	Inserting a reason in Lucidchart	Demonstrating a reason (des)
Ben (M): We need to provide more evidence and provide counterargument to them (hydropower)	Allows students to use a shape to represent a counterargument/counterarguments	Looking at a position, reason, and evidence in Lucidchart	Generating a counterargument/counterarguments against the hydropower team (cla)
…………………………………………			

Table 4 continued

Students' argumentation	Computer's action	Nonverbal action	Codes
Alan (M): Guys, seriously, think! We need more evidence	Shows reasons and evidence that support their position	Looking at a position, reason, and evidence in Lucidchart	Generating an evidence (res)
Quiet for a few minutes………			
Alan (M): Let's look for nuclear plants in Illinois. Let's stay focused	Shows reasons and evidence that support their position	Looking at a position, reason, and evidence in Lucidchart	Generating a evidence (res)
Ben & Celine (M): Wow, they are counterarguing us	Shows a counterargument/counterarguments	Looking at a counterargument in Lucidchart	Demonstrating a counterargument/counterarguments (red)
Alan (M): Hydropower plants can convert up to 90% of the energy to usable electricity, making it very efficient. We should rebut. No, counterargue	Shows a reason and evidence	Looking at the evidence in Lucidchart	Generating a counterargument/counterarguments (cla)
………………………………………			
Alan (M): I don't know if you guys noticed. I put a counterargument. Hydropower plants can harm marine life	Allows students to use a shape to represent a counterargument/counterarguments	Working on Lucidchart	Demonstrating a counterargument (cla)

After assigning the initial codes, the researchers continued to examine if the graph-oriented computer-assisted application's action and code for argumentation skills were associated. The researchers identified patterns that a number of functions of the graph-oriented computer application were associated with types of affordance in the argumentation process as indicated in Table 5. The researchers identified four functions of the computer-assisted application: (a) exhibiting, (b) helping to construct argumentation, (c) eliciting, and (d) acting as an ally. Table 6 indicates how each function relates to type of affordance in the argumentation process.

Additionally, when the researchers coded the data, the researchers identified the distinction between the use of metacognition regulation in the collaborative argumentation process by the female teams and the male teams in both conditions. The researchers assigned the code META to code transcripts of verbal argumentation process. The researchers examined whether there was a similar pattern between female teams' argumentation and male teams' argumentation during the construction of argumentation maps in Lucidchart in the treatment condition. Then, the researchers compared across all teams in both conditions. The metacognition regulation skills included referring to notes and resources (e.g., the science teacher's web site on alternative energy) when planning to argue, articulating each argumentation skills before constructing the map, comparing scientific facts and opinions when constructing the map and wrapping-up, checking grammatical errors for correct text, summarizing at the end of argumentation, and delegating tasks to team members.

Results

Table 7 summarizes the quantitative and qualitative findings.

Research question one

Research question one asked whether students who engaged in collaborative within-gender team argumentation differed in science knowledge from students who engaged in collaborative cross-gender team argumentation. Multivariate analysis of variance (MANOVA) showed no statistically significant effect between the students engaging in within-gender team argumentation ($M = 5.71$, $SD = 1.03$) and students engaging in cross-gender team argumentation ($M = 5.83$, $SD = .61$) on the combined set of science knowledge outcomes (scientific facts, scientific explanations, and valid scientific facts/explanations), $F(3,100) = 1.48$, $p = .22$.

When each gender was considered separately, MANOVA indicated no statistically significant difference in science knowledge between male students engaging in within-gender team argumentation ($M = 5.72$, $SD = 1.16$) and male students engaging in cross-gender argumentation [$M = 5.82$, $SD = .50$; $F(3,47) = 1.93$, $p = .14$]. Similarly, MANOVA indicated no statistically significant difference between female students engaging in within-gender team argumentation ($M = 5.09$, $SD = .89$) and female students engaging in cross-gender team argumentation [$M = 5.82$, $SD = .50$; $F(3,49) = 0.19$, $p = .90$].

MANOVA was chosen because we were interested in the collective set of dependent variables (e.g., scientific facts, scientific explanations, and valid scientific facts/explanations) that conceptually "hang together" and thus are appropriate for MANOVA (Huberty and Olejnik 2006). MANOVA appropriately accounts for covariation among the dependent variables in these situations. The current study and research questions did not involve any

Table 5 Computer application's function and corresponding type of affordance in the argumentation process

Students' argumentation	Function of computer application	Types of affordance in the argumentation process
Alan (M): Nuclear is cool. Our position is nuclear. I need a reason	Helping to construct argumentation	Demonstrating
Alan (M): It can be built anywhere	Helping to construct argumentation	Demonstrating
Ben (M): Yes	Helping to construct argumentation	Demonstrating
Celine (M): Let's look for the stuff there (Lucidchart)	Helping to construct argumentation	Demonstrating
Alan (M): Let's see what they (hydropower team) are putting	Exhibiting	Reading
Alan (M): Do you guys have anything (reason) else?	Eliciting	Responding
Ben (M): They are other reasons	Eliciting	Responding
Alan (M): There are several (nuclear power plants) in Illinois	Helping to construct argumentation	Demonstrating
Celine (M): 6?	Helping to construct argumentation	Demonstrating
Alan (M): Let's get information from iPad	Eliciting	Responding
Alan (M): Unlimited. It creates jobs	Helping to construct argumentation	Demonstrating
Ben and Celine: Yes	Helping to construct argumentation	Demonstrating
Alan (M): Low emission	Helping to construct argumentation	Demonstrating
Ben (M): Maybe require less raw materials	Helping to construct argumentation	Demonstrating
Celine (M): It gets easy transportation	Eliciting	Responding
Alan and Ben (M): What do you mean?	Eliciting	Responding
Alan (M): It's weather reliant	Helping to construct argumentation	Demonstrating
Alan (M): We need evidence to support this (Pointing to a reason in Lucidchart)	Eliciting	Responding
Alan, Ben and Celine review four reasons in Lucidchart and brainstorming	Exhibiting	Metacognition
Alan (M) looking for the cloud bubble. What is the ratio?	Helping to construct argumentation	Demonstrating an evidence
Ben (M): 28 grams of uranium releases as much as energy as 100 tons of coal	Helping to construct argumentation	Demonstrating an evidence
Alan (M): So that supports "Requires little raw material"	Helping to construct argumentation	Demonstrating a reason
Ben (M): We need to provide more evidence and provide counterargument to them (hydropower)	Acting as ally	Claiming
………………………………………		
Alan (M): Guys, seriously, think! We need more evidence	Eliciting	Responding
Quiet for a few minutes………		

Table 5 continued

Students' argumentation	Function of computer application	Types of affordance in the argumentation process
Alan (M): Let's look for nuclear plants in Illinois. Let's stay focused	Eliciting	Responding
Ben & Celine (M): Wow, they are counterarguing us	Exhibiting	Reading
Alan (M): Hydropower plants can convert up to 90% of the energy to usable electricity, making it very efficient. We should rebut. No, counterargue	Acting as ally	Claiming
……………………………………		
Alan (M): I don't know if you guys noticed. I put a counterargument. Hydropower plants can harm marine life	Acting as ally	Claiming

Table 6 Association between the functions of computer application and types of affordance in the argumentation process

Function	Type of affordance
Construct	Demonstrating
Exhibit	Read, metacognition
Elicit	Respond
Ally	Claim (counterargument, rebuttal)

interaction effects. We studied the main effect of grouping type. Regarding assumptions, the group sample sizes were sufficiently large (i.e., greater than 40 per group) and thus the Central Limit Theorem ensures normality of mean values. Similarly, the relative equality of group samples sizes mitigated any potential bias that could be introduced by inequality of variances/covariances (Tabachnick and Fidell 2013). Follow-up tests were not conducted when no main effects were observed, because this would be in no instance an appropriate procedure. Follow-up tests are conducted in MANOVA only when the omnibus test is statistically significant (see Huberty and Olejnik 2006; Tabachnick and Fidell 2013).

Research question two

Research question two asked whether there was a gender-grouping effect on argumentation skill among students who engaged in collaborative argumentation. Multivariate analysis of variance (MANOVA) showed no statistically significant difference between the students engaging in within-gender team argumentation (M = 7.97, SD = 2.64) and students engaging in cross-gender team argumentation (M = 7.65, SD = 2.91) on the combined set of argumentation skill outcomes (reason, evidence, counterargument, and rebuttal), $F(4,99) = 0.62$, $p = .65$.

When each gender was considered separately, MANOVA indicated no statistically significant gender-grouping effect on argumentation skills for females [$F(4,48) = 1.25$, $p = .30$]. However, a marginally significant effect for gender grouping on the combined set of

Table 7 Summary of the findings

Type of data	Research question	Finding
Quantitative data	1	No significant difference
	1a	No significant difference
	2	No significant difference
	2a	Among males, the gender-grouping effect was strongest on rebuttal [$F(1, 49) = 7.34$, $p < .01$]
	3	Among male teams, the gender-grouping effect was strongest on reason [$F(1, 15) = 5.43$, $p < .03$] and rebuttal [$F(1, 15) = 11.38$, $p < .01$]
Qualitative data	4	Summarized in Table 8

Table 8 Comparison among all female and male teams in the within-gender team argumentation and cross-gender team argumentation conditions in the use of argumentation components in argumentation maps in Lucidchart

Condition	Reason	Evidence	Counterargument	Rebuttal
Treatment (within-gender argumentation)	*M*	*M*	*M*	*M*
11 Female Teams	1.55	1.64	2.67	2.00
9 Male Teams	2.67*	1.67	1.78	0.44*
Control (cross-gender argumentation)				
8 female teams	1.25	1.38	1.25	1.25
8 male teams	1.25*	1.25	1.75	1.50*

* $p < .05$

outcomes was apparent for males [$F(4,46) = 2.54$, $p = .05$]. Examination of the canonical loadings (i.e., structure coefficients; -0.43, 0.11, 0.15, 0.82, for reason, evidence, counterargument, and rebuttal, respectively) indicated that the gender-grouping effect was strongest for rebuttal. Univariate ANOVA analyses also affirmed a statistically significant gender-grouping effect on rebuttal [$F(1,49) = 7.34$, $p < .01$], with a moderate-to-large effect size ($\eta^2 = .13$). Here, the mean rebuttal score among male students engaging in cross-gender team argumentation ($M = 1.68$, $SD = 1.25$) was higher than the mean score for students engaging in within-gender team argumentation ($M = 0.83$, $SD = 1.00$).

Research question three

Research question three asked whether there are differences in argumentation maps in Lucidchart (as measured by reason, evidence, counterargument, and rebuttal) between female teams (or male teams) engaging in within-gender team argumentation (treatment condition) and female teams (or male teams) engaging in cross-gender team argumentation. As indicated in Table 8, when the set of four argumentation maps subscale scores (reason, evidence, counterargument, and rebuttal) were considered as outcome variables, MANOVA showed a statistically significant effect for gender grouping among male teams [$F(4, 12) = 3.74$, $p = .03$]. No statistically significant effect was observed for female teams [$F(4, 14) = 1.23$, $p = .34$]. Examination of the canonical loadings (-0.54, -0.28, -0.01, and 0.78 for reason, evidence, counterargument, and rebuttal, respectively) indicated that, among male teams, the gender-grouping effect was strongest for reason and rebuttal. Follow-up univariate ANOVAs on the four subscale scores showed a statistically significant gender-grouping effect on both of these outcomes [$F(1, 15) = 5.43$, $p = .03$ for reason; and $F(1, 15) = 11.38$, $p < .01$ for rebuttal], while the gender effect for the other univariate outcomes was not statistically significant. Effect sizes for the effect of gender on reason and rebuttal were large ($\eta^2 = .27$ for reason and $\eta^2 = .15$ for rebuttal). Here, the mean reason score among male teams for the cross-gender team argumentation condition ($M = 1.25$) was lower than the mean score for the within-gender team argumentation condition ($M = 2.67$). However, the mean rebuttal score among male teams for the cross-gender team argumentation condition ($M = 1.5$) was higher than the mean score for male teams in the within-gender team argumentation condition ($M = 0.44$).

Research question four

Question four asked how the GOCAA supported female and male teams' development of argumentation skills in different types of gender groupings. As indicated in Table 9, a number of patterns were identified. First, the application helped all teams in both conditions to construct their argumentation by demonstrating their position, reasons, evidence, counterargument or rebuttals on Lucidchart.

Second, the application is an external representation that exhibits argumentation process. All teams in both conditions made sense of the process by reading the external representation. They could read position, reasons, and evidence provided either by their team or the other team. Moreover, all female teams in both conditions looked at the application and showed a number of metacognition regulation behaviors. For example, the female students tended to look at Lucidchart and make statement such as, "Let's take a look at notes and find disadvantages and advantages of (type of energy)" and "Let's check out Dr. V's (science teacher) website for more information." The female students tended to verbally explain what five argumentation skills are with their team members and ensured everyone had the same understanding when they inserted different shapes into the work area in Lucidchart. For example, they would say, "Rebuttal is to rebut whatever they said about us." or "Let's stick to facts not opinion." They tended to share the workload during the collaborative argumentation process. They would say, "We can take turn typing." Or team members provided ideas to the

Table 9 The number of female and male teams demonstrating application functions and types of affordances in the within-gender team argumentation and cross-gender team argumentation conditions

Functions/type affordance	Treatment condition (within-gender argumentation)		Control condition (cross-gender argumentation)	
	11 female teams	9 male teams	8 female teams	8 male teams
Constructing argumentation/demonstrating	11	9	8	8
Exhibiting/reading	11	9	8	8
Exhibiting/regulating metacognition	11	1	8	1
Eliciting/responding (reason, evidence, rebuttal)	11	9	8	8
Acting as ally/claiming (counterargument against the Tawfik other team)	11	9	8	8
Acting as ally/claiming (rebuttal)	11	2	8	8

team member who typed. When they wrapped up the activity, the female students tended to summarize their argumentation process and looked for grammatical errors on the argumentation map in Lucidchart. However, metacognition regulation is only evident in one out of nine male teams in the treatment condition and one out of eight male teams in the control condition. The male students tended to leave their notes behind. When the teacher inquired where their notes are, they would say, “They are all in my brain.” The male students also tended to confuse with the definition of each argumentation skill until the researchers provided clarification. Some teams even forgot what topic they should argue. In most of the male teams, only one student was in charge of typing and thinking, while other team members simply sat there or walked around.

Third, the application elicits student response. Looking at the external representation of their positions, all teams in both conditions responded by providing reasons and evidence to support their position. Fourth, when acting as an ally, the application is used by students to support their efforts to make a case by claiming a counterargument or rebuttal. Specifically, all teams in both conditions claimed a counterargument or counterarguments against their corresponding team to make their case. However, with the exception of seven male teams in the treatment condition, all teams in both conditions claimed a rebuttal or rebuttals against a counterargument or counterarguments provided by their corresponding team.

In the collaborative argumentation process, all teams in both conditions negotiated which shape should be used for the corresponding argumentation skill. In the treatment condition, nine female (out of 11) and two male teams (out of eight) did a good job in aligning the shape and the corresponding argumentation skill in the argumentation maps in Lucidchart. In the control condition, seven female (out of eight) and seven male teams (out of eight) did a good job in aligning the shape and the corresponding argumentation skill in the argumentation maps in Lucidchart. In the control condition, the female and male teams tended to look at the argumentation maps in Lucidchart that the other team did and learned from it. For example, when the students brainstormed their argumentation, they would say, “Let’s see what they have here.” or “Let’s check out what they did.”

Discussion

In this study, the findings are inconsistent with previous research (Asterhan et al. 2012; Ding et al. 2009, 2011; Sullivan et al. 2015; Zhan et al. 2015) in terms of the effect of different types of gender groupings on students’ content knowledge in computer-assisted collaborative learning. Overall, the previous research suggested that female students tend to perform better in content knowledge in homogeneous-gender groupings or same-gender pairings than female students in the mixed-gender groupings or pairs. The present study found no significant difference in science knowledge between students (either female or male) engaging in within-gender team argumentation (treatment condition) and students (either female or male) engaging in cross-gender team argumentation (control condition). The findings might be attributed to a number of reasons. First, the students in some of the previous studies were intentionally separated from their team member(s) by placing each of them in different rooms or in different geographical areas when they collaborated with their team member(s) in a computer-supported collaborative learning environment but were not given sufficient opportunities to engage in verbal argumentation in a synchronous way. However, in the present study, all the students were engaged in verbal collaborative argumentation with the members of their own homogeneous-gender team both before and while they argued with either another homogeneous-gender team (using either within-gender argumentation or cross-gender argumentation) with the support of the GOCAA. The verbal collaborative argumentation facilitated a common tempo for all the team members to share their science knowledge, which could avoid the issue of the varied tempos for participants in computer-supported collaborative learning due to lack of shared contexts (Weinberger et al. 2010). Second, verbal collaborative argumentation would put female students at great advantage because they tend to communicate ideas better in verbal form (Tarim and Kyratzis 2012), which allows each member to contribute their science knowledge. Third, in previous studies (e.g., Ding et al. 2011), the students collaborated with their team member(s) via text-based computer-supported learning environments. It is difficult for students to track and monitor texts in a chat room because the text threads may become lengthy. However, in this study, the students were asked to make their science knowledge visible using different shapes that represent different argumentation skills in the GOCAA. The GOCAA became a space to share science knowledge, and each student was able to track, monitor, and use the shared science knowledge in this space (Lajoie and Derry 2013). Fourth, in this study, the students were asked to use visual elements (e.g., rectangles to indicate reasons) to make their ideas visible on the GOCAA. Although the GOCAA became a tool to represent science knowledge visually and could support male students' visual mode of knowledge representation, the GOCAA can support female students' verbal mode of knowledge representation by allowing them to type in texts in the shapes in GOCAA, which might reduce the opportunities for females to develop divergent pattern of knowledge elaboration (Buck et al. 2012).

This study found no significant difference in argumentation skills between female students engaging in within-gender team argumentation (treatment condition) and female students engaging in cross-gender team argumentation (control condition). However, this study found a marginally significant difference in argumentation skills between male students engaging in within-gender team argumentation (treatment condition) and male students engaging in cross-gender team argumentation (control condition). Male students in engaging in within-gender team argumentation (treatment condition) performed worse than male students engaging in cross-gender argumentation (control condition) in argumentation skills on individual essays, particularly in rebuttal skills. This gender difference in the effect of gender grouping on argumentation skills could be attributed to a number of reasons. First, the collaborative argumentation process involved extensive planning and communication among team members. Research (Korobov 2013; Stokoe 2004; Tarim and Kyratzis 2012) has shown that female students are better than male students at planning and communicating and may engage in more discussion during computer-supported collaborative learning than male teams, ultimately leading to a better outcome, such as better argumentation skills. In this study, the female teams, regardless of the type of gender grouping, demonstrated balanced participation in the collaborative argumentation process, an effect that could be reflected by the similar number of representations of each argumentation skill in their argumentation maps in Lucidchart. The female teams also demonstrated a similar pattern in terms of types of affordances in both conditions. On the contrary, male teams engaging in within-gender team argumentation (treatment condition) demonstrated unbalanced participation in the collaborative argumentation process, an effect that could be reflected by the statistically significant difference between reason and rebuttal components in their argumentation maps in Lucidchart. These observations suggested that male students tended to lack planning and communication skills, which can lead to unbalanced involvement and participation in the collaborative argumentation process.

Second, this gender difference in the effect of gender grouping on argumentation skills could also be attributed to the uses of metacognition regulation skills. In this study, the exhibiting function of the GOCAA afforded the female teams opportunities to develop metacognition regulation skills in both conditions. On the contrary, these behaviors were evident in one out of nine male teams engaging in within-gender team argumentation (treatment condition) and one out of eight male teams engaging in cross-gender argumentation (control condition). The male teams, regardless of the type of gender grouping, showed poor use of metacognition regulation skills. In this study, the findings are similar to previous findings (Asterhan et al. 2012; Caspi et al. 2008; Ma and Yuen 2011; Prinsen et al. 2007). The poor use of metacognition regulation skills by male teams might result in the poor quality of the degree of participation, kind of participation, and experience of participation in a computer-assisted collaborative learning environment, leading to poor learning outcome, that is, argumentation skills in this study.

However, this study also showed that the metacognition regulation skills of the male teams could be enhanced by working with the female teams. In the control condition (cross-gender team argumentation), the female and male teams tended to look at the argumentation maps in Lucidchart that the other team had constructed, and learned from them. This is an example of the spreading effect (Kim et al. 2007). In cross-gender team argumentation (control condition), the male teams were aware that the other team was a female team and seemed to pay more close attention to the structure of argumentation maps developed by the corresponding team and learned from it. It is evident that male teams in the cross-gender team argumentation performed better in the alignment between shapes/arrows and corresponding argumentation skill than male teams engaging in within-gender team argumentation.

In science education, earlier research about supporting students' scientific argumentation skills has focused on the written form such as the Science Writing Heuristic (Cavagnetto et al. 2010). Recently, researchers have taken different approaches to develop students' scientific argumentation skills such as engaging students in argumentation talk and have portrayed argumentation as a social process of constructing, supporting, and critiquing the claims for the purpose of developing shared knowledge (Berland and Reiser 2009; Ryu and Sandoval 2012). Along with this line of research, the scholars also caution that more work is needed to understand types of support needed to engage students in this social process. This study advances our knowledge of the importance of gender differences in communication styles and cognitive abilities when involving students of different gender groupings in the collaborative argumentation process. As indicated in this study, using a GOCAA to support collaborative argumentation process has potential in mediating the differences of female and male students' development of science knowledge. Yet, researchers need to explore ways (e.g., prompts in the GOCAA, teacher guidance) to address male students' unbalanced participation and lack of metacognition skills in the collaborative argumentation process.

Limitations and future studies

There are a number of limitations in this study. The findings showed that the male teams tended to use metacognitive regulation skills to a lesser extent than their female teams to regulate their collaborative argumentation process supported by a GOCAA. One limitation is that metacognitive regulation is one type of affordance emerging from our qualitative analysis. Further studies are needed to explore what causes male students to fail to use their metacognitive skills and the affordances of the GOCAA, and to assess metacognitive regulation in a quantitative manner that allows for statistical analysis. Additionally, conducting a qualitative research study, adopting in-depth interviews with the male students, could provide insight into number of questions. Do male students have difficulty in impulse control? Do they find other male students' competition distracting? Are they aware that they are failing to be strategic in their work? When they work with the female teams, do they notice the difference in the quality of argumentation maps? Further studies are also needed to examine if there are any interventions such as a teacher's guidance (Hsu et al. 2015) and the GOCAA's built-in features (Quintana et al. 2004) that could reinforce the female teams' metacognition regulation skills and improve the male teams' metacognition regulation skills and how. For example, if the teacher provides students with map-rating and argumentation scoring rubrics, would they improve their performance? The other limitation is that participants' science learning competencies were not controlled. Further study could be conducted that incorporates one or more pretests to control for possible variation due to participants' prior science learning experiences, or their science knowledge on a particular topic.

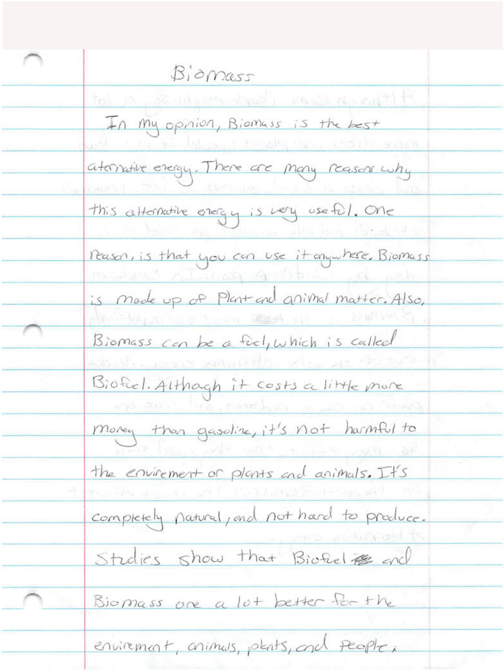
This study was conducted with students from middle-class families. The school was equipped with rich technology and abundant opportunity for student experience with project-based science learning. Further studies are needed to examine how classroom variables such as socio-economic status, level of technology resources, lecture-based science teaching, cultural background (Vatrapu 2008) may affect the outcomes of this study. This study was conducted within one semester for one science topic. Longitudinal studies are needed to examine how different types of gender groupings affect female and male students in science knowledge and argumentation skills. Additionally, this study is limited to the context of middle-level students, and further studies are needed that examine students in different settings, such as a college-level setting.

Implications and conclusion

This study proposed that engaging middle-level female students in scientific argumentation in computer-supported collaborative learning environments could enhance their interests in science. It is important for instructional designers and educators to consider gender differences in cognitive abilities and communication styles as well as different types of gender groupings in orchestrating such learning environments (Dickey 2014). This study concluded that, when integrating verbal collaborative argumentation activity to bring all team members to the same tempo in a GOCAA learning environment and using a GOCAA to support the argumentation, female students engaging in cross-gender team argumentation condition do not perform worse than female students engaging in within-gender team argumentation. Further, these female students engaging in cross-gender team argumentation might be less likely to employ divergent patterns of knowledge elaboration due to the differences of communication styles and cognitive abilities between female and male students.

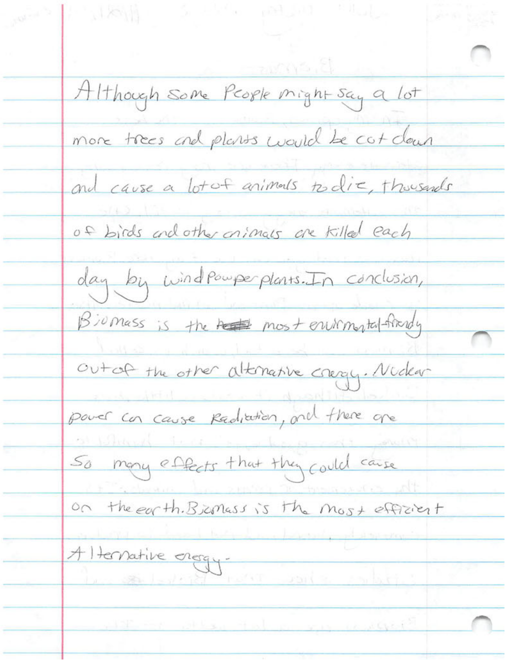
However, it was a disadvantage for the male students to engage in the process with the male students in the within-gender team argumentation. Instructional designers or educators could use this observation to plan computer-assisted collaborative learning in different stages. For example, it may be beneficial first to engage students in the same-gender argumentation and then to provide them with the opportunity for cross-gender argumentation. Practitioners need to be cautious about the use of different types of gender grouping strategies, so that neither male nor female students are placed at disadvantage. In this study, cross-gender team argumentation appears to benefit both female and male students.

Appendix 1: Student sample post essay



Biomass

In my opinion, Biomass is the best alternative energy. There are many reasons why this alternative energy is very usefol. One reason, is that you can use it anywhere. Biomass is made up of Plant and animal matter. Also, Biomass can be a fuel, which is called Biofuel. Although it costs a little more money than gasoline, it's not harmful to the envirement or plants and animals. It's completely natural, and not hard to produce. Studies show that Biofuel ~~is~~ and Biomass are a lot better for the envirement, animals, plants, and people.



Although some People might say a lot more trees and plants would be cut down and cause a lot of animals to die, thousands of birds and other animals are killed each day by wind Powper plants. In conclusion, Biomass is the ~~best~~ most enviormental-friendly out of the other alternative energy. Nuclear power can cause Radiation, and there are so many effects that they could cause on the earth. Biomass is the most efficient Alternative energy.

Appendix 2: Science knowledge rubric

Criteria	0 points	1 point	2 points	3 points
Scientific facts presented	No facts presented	1 fact presented	2 facts presented	3 or more facts presented
Scientific explanation stated to support facts	No explanation	1 statement of explanation	2 statements of explanation	3 statements of explanation
Scientific validity of facts/explanation	3 or more facts/explanation statements are false	2 facts/explanation statements are false	1 fact/explanation statement is false	All statements are valid

Appendix 3: Correct logic for scoring argumentation skills in individual essay

Example 1:

(Indicate a Position)–(Indicate Reasons)–(Indicate Evidence)–(Indicate Counterarguments)–(Indicate Rebuttals)

Example 2:

(Indicate a Position)–(Indicate Reasons)–(Indicate Evidence)–(Indicate Counterargument)–(Indicate Rebuttal))–(Indicate Counterargument)–(Indicate Rebuttal)

References

Abnett, C., Stanton, D., Neale, H., & O'Malley, C. (2001). The effect of multiple input devices on collaboration and gender issues. In *European perspectives on computer-supported collaborative learning (EuroCSCL) 2001* (pp. 29–36). Maastricht: University of Bath. Retrieved from http://opus.bath.ac.uk/9674/.

Andriessen, J. (2006). Arguing to learn. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 443–460). New York: Cambridge University Press.

Andriessen, J., Baker, M., & Suthers, D. (2003). Argumentation, computer support, and the educational context of confronting cognitions. *Arguing to learn* (pp. 1–25). Dordrect: Springer.

Asterhan, C. S. C., Schwarz, B. B., & Gil, J. (2012). Small-group, computer-mediated argumentation in middle-school classrooms: The effects of gender and different types of online teacher guidance. *British Journal of Educational Psychology, 82,* 375–397.

Baram-Tsabari, A., & Yarden, A. (2011). Quantifying the gender gap in science interests. *International Journal of Science and Mathematics Education, 9*(3), 523–550. https://doi.org/10.1007/s10763-010-9194-7.

Berland, L. K., & Reiser, B. J. (2009). Making sense of argumentation and explanation. *Science Education, 93*(1), 26–55.

Bransford, J., & Johnson, M. (1973). Consideration of some problems of comprehension. In W. Chase (Ed.), *Visual information processing* (pp. 383–483). New York: Academic Press.

Brotman, J. S., & Moore, F. M. (2008). Girls and science: A review of four themes in the science education literature. *Journal of Research in Science Teaching, 45*(9), 971–1002. https://doi.org/10.1002/tea.20241.

Buck, G. A., Beeman-Cadwallader, N. M., & Trauth-Nare, A. E. (2012). Keeping the girls visible in K-12 science education efforts: A feminist case study on problem-based learning. *Journal of Women and Minorities in Science and Engineering, 18*(2), 153–178. https://doi.org/10.1615/JWomenMinorScienEng.2012002317.

Buffery, A., & Gray, J. (1972). Sex differences in the development of spatial and linguistic skills. In C. Ounsted & D. C. Taylor (Eds.), *Gender differences, their ontogeny and significance* (pp. 123–158). Edinburgh: Churchill Livingstone.

Caplan, P. J., Crawford, M., Hyde, J. S., & Richardson, J. T. E. (1997). *Gender differences in human cognition*. New York: Oxford University Press.

Carlone, H. B., Scott, C. M., & Lowder, C. (2014). Becoming (less) scientific: A longitudinal study of students' identity work from elementary to middle school science. *Journal of Research in Science Teaching, 51*(7), 836–869.

Carr, C. S. (2003). Visualizing argumentation: Software tools for collaborative and educational sense making. In P. A. Kirschner, S. J. Buckingham Shum, & C. S. Arr (Eds.), *Using computer supported argument visualization to teach legal argumentation* (pp. 75–96). London: Springer.

Caspi, A., Chajut, E., & Saporta, K. (2008). Participation in class and in online discussions: Gender differences. *Computers & Education, 50*(3), 718–724. https://doi.org/10.1016/j.compedu.2006.08.003.

Cavagnetto, A., Hand, B. M., & Norton-Meier, L. (2010). The nature of elementary student science discourse in the context of the science writing heuristic approach. *International Journal of Science Education, 32*(4), 427–449.

Chen, X. (2013). *STEM attrition: College students' paths into and out of STEM fields* (NCES 2014-001). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Cho, K. L., & Jonassen, D. H. (2003). The effects of argumentation scaffolds on argumentation and problem solving. *Educational Technology Research and Development, 50*(3), 5–22.

Chung, C.-W., Lee, C.-C., & Liu, C.-C. (2013). Investigating face-to-face peer interaction patterns in a collaborative Web discovery task: the benefits of a shared display. *Journal of Computer Assisted Learning, 29,* 188–206.

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. Los Angeles, CA: Sage.

Crowell, A., & Kuhn, D. (2014). Developing dialogic argumentation skills: A three-year intervention study. *Journal of Cognition and Development, 15*(2), 363–381. https://doi.org/10.1080/15248372.2012.725187.

Dickey, M. W. (2014). *Gender-specific instructional strategies and student achievement in 5th grade classrooms.* Doctoral Dissertation. Retrieved from http://scholarcommons.sc.edu/etd/2624.

Ding, N., Bosker, R. J., & Harskamp, E. G. (2009). How gender composition influences individual knowledge elaboration in CSCL. In *Proceedings of the 9th international conference on Computer supported collaborative learning* (Vol. 1, pp. 173–177). Rhodes, Greece: International Society of the Learning Sciences. Retrieved from http://dl.acm.org/citation.cfm?id=1600053.1600079.

Ding, N., Bosker, R. J., & Harskamp, E. G. (2011). Exploring gender and gender pairing in the knowledge elaboration processes of students using computer-supported collaborative learning. *Computers & Education, 56*(2), 325–336.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education, 84*(3), 287–312.

Dwyer, C. P., Hogan, M. J., & Stewart, I. (2012a). An evaluation of argument mapping as a method of enhancing critical thinking performance in e-learning environments. *Metacognition and Learning, 7*(3), 219–244. https://doi.org/10.1007/s11409-012-9092-1.

Dwyer, C. P., Hogan, M. J., & Stewart, I. (2012b). An evaluation of argument mapping as a method of enhancing critical thinking performance in e-learning environments. *Metacognition and Learning, 7*(3), 219–244. https://doi.org/10.1007/s11409-012-9092-1.

Easterday, M. W., Aleven, V., Scheines, R., & Carver, S. M. (2009). Constructing causal diagrams to learn deliberation. *International Journal of Artificial Intelligence in Education, 19*(4), 425–445.

Evagorou, M., & Osborne, J. (2013). Exploring young students' collaborative argumentation within a socioscientific issue. *Journal of Research in Science Teaching, 50*(2), 209–237. https://doi.org/10.1002/tea.21076.

Fairweather, H. (1976). Sex differences in cognition. *Cognition, 4,* 231–280.

George, R. (2006). A cross-domain analysis of change in students' attitudes toward science and attitudes about the utility of science. *International Journal of Science Education, 28,* 571–589.

Halpern, D. F. (1992). *Sex differences in cognitive abilities* (2nd ed.). Hillsdale, NJ: Erlbaum.

Halpern, D., Aronson, J., Reimer, N., Simpkins, S., Star, J., & Wentzel, K. (2007). *Encouraging girls in math and science* (NCER 2007-2003). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ncer.ed.gov.

Hoffmann, L. (2002). Promoting girls' interest and achievement in physics classes for beginners. *Learning and instruction, 12*(4), 447–465.

Hogan, K., Nastasi, B. K., & Pressley, M. (1999). Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognition and Instruction, 17*(4), 379–432.
Hsu, P.-S., Van Dyke, M., Chen, Y., & Smith, T. J. (2015). The effect of a graph-oriented computer-assisted project-based learning environment on argumentation skills. *Journal of Computer Assisted Learning, 31*(1), 32–58.
Hsu, P.-S., Van Dyke, M., Chen, Y., & Smith, T. J. (2016). A cross-cultural study of the effect of a project-based learning environment that incorporates a graph-oriented, computer-assisted application on middle school students' science knowledge and scientific argumentation. *Journal of Computer Assisted Learning, 32*(1), 51–76.
Huberty, C. J., & Olejnik, S. (2006). *Applied MANOVA and discriminant analysis* (Vol. 498). New York: Wiley.
Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin, 104,* 53–69.
Iordanou, K. (2010). Developing argument skills across scientific and social domains. *Journal of Cognition and Development, 11*(3), 293–327. https://doi.org/10.1080/15248372.2010.485335.
Jonassen, D. H., & Kim, B. (2010). Arguing to learn and learning to argue: Design justifications and guidelines. *Educational Technology Research and Development, 58,* 439–457.
Kelly, A. (1978). *Girls and science: An international study of sex differences in school science achievement*. IEA Monograph Studies, No. 9. Stockholm, Sweden: Almqvistwiksell International.
Kelly, G. J., & Crawford, T. (1996). Student's interaction with computer representations: Analysis of discourse in laboratory groups. *Journal of Research in Science Teaching, 33*(7), 693–707. https://doi.org/10.1002/(SICI)1098-2736(199609)33:7<693:AID-TEA1>3.0.CO;2-I.
Kiili, C. (2012). Argument graph as a tool for promoting collaborative online reading. *Journal of Computer Assisted Learning, 29*(3), 248–259. https://doi.org/10.1111/j.1365-2729.2012.00492.x.
Kim, I.-H., Anderson, R. C., Nguyen-Jahiel, K., & Archodidou, A. (2007). Discourse patterns during children's collaborative online discussions. *Journal of the Learning Sciences, 16*(3), 333–370. https://doi.org/10.1080/10508400701413419.
Korobov, N. (2013). Positioning identities: A discursive approach to the negotiation of gendered categories. *Narrative Inquiry, 23*(1), 111–131. https://doi.org/10.1075/ni.23.1.06kor.
Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. *Science Education, 77*(3), 319–337. https://doi.org/10.1002/sce.3730770306.
Kuhn, D. (2011). Teaching and learning science as argument. *Science Education, 94*(5), 810–824.
Kuhn, D. (2015). Thinking together and alone. *Educational Researcher, 44*(1), 46–53. https://doi.org/10.3102/0013189X15569530.
Kuhn, D., Goh, W., Iordanou, K., & Shaenfield, D. (2008). Arguing on the computer: A microgenetic study of developing argument skills in a computer-supported environment. *Child Development, 79*(5), 1310–1328. https://doi.org/10.1111/j.1467-8624.2008.01190.x.
Kuhn, D., & Udell, W. (2003). The development of argument skills. *Child Development, 74*(5), 1245–1260. https://doi.org/10.1111/1467-8624.00605.
Kuhn, D., Wang, Y., & Li, H. (2010). Why argue? Developing understanding of the purposes and values of argumentive discourse. *Discourse Processes, 48*(1), 26–49. https://doi.org/10.1080/01638531003653344.
Lajoie, S. P., & Derry, S. J. (2013). *Computers as cognitive tools*. New York: Routledge.
Leaper, C. (1991). Influence and involvement in children's discourse: Age, gender, and partner effects. *Child Development, 62,* 797–811.
Ma, W. W. K., & Yuen, A. H. K. (2011). Gender differences of knowledge sharing in online learning environment. In R. Kwan, J. Fong, L. Kwok, & J. Lam (Eds.), *Hybrid Learning* (pp. 116–128). Berlin: Springer. https://doi.org/10.1007/978-3-642-22763-9_11.
Maltz, D. N., & Borker, R. (1982). A cultural approach to male-female miscommunication. In J. J. Gumpertz (Ed.), *Language and social identity*. Cambridge: Cambridge University Press.
Manz, E. (2015). Representing student argumentation as functionally emergent from scientific activity. *Review of Educational Research, 85*(4), 553–590.
National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Committee on a Conceptual Framework for New K-12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.
Pea, R. D. (1985). Beyond amplification: Using the computer to reorganize mental functioning. *Educational Psychologist, 20*(4), 167–182. https://doi.org/10.1207/s15326985ep2004_2.
Prinsen, F. R., Volman, M. L. L., & Terwel, J. (2007). Gender-related differences in computer-mediated communication and computer-supported collaborative learning. *Journal of Computer Assisted Learning, 23*(5), 393–409. https://doi.org/10.1111/j.1365-2729.2007.00224.x.

Quinn, D. M., & Cooc, N. (2015). Science achievement gaps by gender and race/ethnicity in elementary and middle school trends and predictors. *Educational Researcher, 44*(6), 336–346.

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., et al. (2004). A scaffolding design framework for software to support science inquiry. *The Journal of the Learning Sciences, 13*(3), 337–386.

Robinson, A., Pérez-Quiñones, M. A., & Scales, G. (2016). African-American middle school girls: Influences on attitudes toward computer science. *Computing in Science & Engineering, 18*(3), 14–23.

Ryu, S., & Sandoval, W. A. (2012). Improvements to elementary children's epistemic understanding from sustained argumentation. *Science Education, 96*(3), 488–526.

Sampson, V., Enderle, P. J., & Walker, J. P. (2012). The development and validation of the assessment of scientific argumentation in the classroom (ASAC) observation protocol: A tool for evaluating how students participate in scientific argumentation. *Perspectives on scientific argumentation* (pp. 235–264). Dordrecht: Springer.

Scheuer, O., Loll, F., Pinkwart, N., & McLaren, B. M. (2010). Computer-supported argumentation: A review of the state of the art. *International Journal of Computer-Supported Collaborative Learning, 5*(1), 43–102. https://doi.org/10.1007/s11412-009-9080-x.

Schwarz, B. B., & Glassner, A. (2007). The role of floor control and of ontology in argumentative activities with discussion-based tools. *International Journal of Computer-Supported Collaborative Learning, 2*(4), 449–478. https://doi.org/10.1007/s11412-007-9024-2.

Stokoe, E. H. (2004). Gender and discourse, gender and categorization: Current developments in language and gender research. *Qualitative Research in Psychology, 1*(2), 107–129. https://doi.org/10.1191/1478088704qp007oa.

Strijbos, J.-W. (2011). Assessment of (computer-supported) collaborative learning. *IEEE Transactions on Learning Technologies, 4*(1), 59–73. https://doi.org/10.1109/TLT.2010.37.

Sullivan, F. R., Kapur, M., Madden, S., & Shipe, S. (2015). Exploring the role of "gendered" discourse styles in online science discussions. *International Journal of Science Education, 37*(3), 484–504. https://doi.org/10.1080/09500693.2014.994113.

Suthers, D. D., & Hundhausen, C. D. (2003). An experimental study of the effects of representational guidance on collaborative learning processes. *Journal of the Learning Sciences, 12*(2), 183–218. https://doi.org/10.1207/S15327809JLS1202_2.

Suthers, D. D., Vatrapu, R., Medina, R., Joseph, S., & Dwyer, N. (2008). Beyond threaded discussion: Representational guidance in asynchronous collaborative learning environments. *Computers & Education, 50*(4), 1103–1127. https://doi.org/10.1016/j.compedu.2006.10.007.

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

Tarim, S. D., & Kyratzis, A. (2012). Challenging and orienting to monolingual school norms in Turkish American children's peer disputes and classroom negotiations at a U.S. Turkish Saturday school. *Sociological Studies of Children and Youth, 15,* 193–220. https://doi.org/10.1108/S1537-4661(2012)0000015012.

Vatrapu, R. K. (2008). Cultural considerations in computer supported collaborative learning. *Research and Practice in Technology Enhanced Learning, 3*(2), 159–201. https://doi.org/10.1142/S1793206808000501.

Velayutham, S., Aldridge, J. M., & Fraser, B. (2012). Gender differences in student motivation and self-regulation in science learning: A multi-group structural equation modeling analysis. *International Journal of Science and Mathematics Education, 10*(6), 1347–1368.

Walton, D. N. (1996). *Argumentation schemes for presumptive reasoning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Weinberger, A., Stegmann, K., & Fischer, F. (2010). Learning to argue online: Scripted groups surpass individuals (unscripted groups do not). *Computers in Human Behavior, 26*(4), 506–515. https://doi.org/10.1016/j.chb.2009.08.007.

Welch, A. G., Cakir, M., Peterson, C. M., & Ray, C. M. (2014). The relationship between gender and classroom environment in Turkish science classrooms. *Educational Research and Reviews, 9*(20), 893.

Wood, J. T. (1996). *Gendered lives: Communication, gender and culture* (2nd ed.). Belmont, CA: Wadsworth.

Zhan, Z., Fong, P. S. W., Mei, H., & Liang, T. (2015). Effects of gender grouping on students' group performance, individual achievements and attitudes in computer-supported collaborative learning. *Computers in Human Behavior, 48,* 587–596. https://doi.org/10.1016/j.chb.2015.02.038.

Pi-Sui Hsu is an Associate Professor of Department of Educational Technology, Research and Assessment (ETRA) at Northern Illinois University. Her research interests focus on technology integration and scientific argumentation.

Margot Van Dyke is a middle school science teacher. Her research interests focus on instructional design and integration of technology in science teaching and learning.

Thomas Smith is a Professor of Department of Educational Technology, Research and Assessment (ETRA) at Northern Illinois University. His research interests focus on quantitative research methodology and large-scale data analysis.

Chee-Kit Looi is Professor of Education at the National Institute of Education, Nanyang Technological University, Singapore. His research interests focus on computer-supported collaborative learning, mobile learning, and computational thinking.