JOURNAL OF RESEARCH IN SCIENCE TEACHING

Argumentation and Participation Patterns in General Chemistry Peer-Led Sessions

Ushiri Kulatunga,[1] Richard S. Moog,[2] and Jennifer E. Lewis[1]

[1]Department of Chemistry, University of South Florida, 4202 E. Fowler Ave. CHE205, Tampa, Florida 33620

[2]Department of Chemistry, Franklin & Marshall College, P.O. Box 3003, Lancaster, Pennsylvania 17604-3003

Received 1 December 2011; Accepted 4 August 2013

Abstract: This article focuses on the use of Toulmin's argumentation scheme to investigate the characteristics of student group argumentation in Peer-Led Guided Inquiry sessions for a General Chemistry I course. A coding scheme based on Toulmin's [Toulmin [1958] The uses of argument. Cambridge: Cambridge University Press] argumentation model was used for identifying arguments during group work without instructor intervention. A modification of the framework developed by Erduran et al. [Erduran, Simon, & Osborne [2004] Science Education, 88(6), 915–933] for characterizing arguments was employed that considered both the strength of the argument and whether an argument contained contributions from one or more than one student. Data were collected by video recording weekly peer-led sessions with a focus on two small groups. Analysis of this video data with the coding scheme and the framework revealed that students were mostly engaged in co-constructed arguments, with more than one student providing evidence and reasoning during group activities. Students often supported their claims with data and warrants but rarely offered backings. That is, they supported their answers with evidence and reasoning but did not often elaborate on their reasoning or further validate their explanations. However, the percentage of arguments containing backings increased when arguments contained contributions from more than one student rather than being presented by one individual. Another significant finding is that students were able to resolve wrong claims through argumentation without peer leader intervention, an indication of independent learning. © 2013 Wiley Periodicals, Inc. J Res Sci Teach

Keywords: argumentation; cooperative grouping; chemistry; first-year undergraduate

The effectiveness of cooperative group learning in post-secondary education has been the subject of much study (Johnson, Johnson, & Smith, 1998; Springer, Stanne, & Donovan, 1999). One common approach to group learning is Peer-Led Team Learning, or PLTL (Gosser, Kampmeier, & Varma-Nelson, 2010; Quitadamo, Brahler, & Crouch, 2009; Sperry & Tedford, 2008), in which students work in groups on problems with a peer leader, a student who has previously taken the same course and performed well. Peer leaders also undergo appropriate training and receive ongoing support throughout the semester to ensure they are prepared for working with students. Previous studies have shown that peer-led sessions can enhance student performance in chemistry. A comparison of participants in peer-led sessions with non-participants showed that students who participated in the peer-led sessions earned higher final grades in

---

Correspondence to: J. E. Lewis; E-mail: jennifer@usf.edu

DOI 10.1002/tea.21107

general chemistry (Drane, Smith, Light, Pinto, & Swarat, 2005) and in organic chemistry (Tien, Roth, & Kampmeier, 2002). Studies also have shown that peer-led sessions have helped students' cognitive and personal growth (Micari, Streitwieser, & Light, 2006). Recently, we have investigated the use of a slightly modified peer-led group-learning approach associated with a large lecture section in general chemistry (Lewis & Lewis, 2005). Peer-Led Guided Inquiry (PLGI) is a cooperative small-group inquiry learning method in which students co-construct targeted chemistry concepts through specially designed paper-and-pencil activities. The major differences with the PLTL approach are that the student groups are smaller (three or four students) with assigned individual roles, there is only one peer leader responsible for five or six groups in a room, and the materials that the students are working with are specially designed to develop, rather than reinforce, important course concepts. In our previous work, we have investigated questions of effectiveness and equity in this group-learning context (Lewis and Lewis, 2005, 2008). Although we have seen that the approach has merit in terms of improving student exam scores in general chemistry, the sources of this improvement are not completely clear. In particular, one significant issue that bears examining is the nature of discourse within the student groups, including the relationship between this discourse and student performance. Are all of the students involved in the discussions, or does one person tend to dominate and "provide the answers"? What relationship, if any, can be seen between student participation in this discourse and overall student performance? Does the discourse contain the components of scientific argumentation, including the presentation of evidence rather than simply the generation of responses? Do the groups generate incorrect answers or inappropriate conceptual constructions that are not corrected? In this work, we examine these questions through the lens of the Toulmin argumentation scheme (Toulmin, 1958). In order to do so, we modify an analytical framework from the work of Erduran, Simon, and Osborne (2004) to address our specific research questions.

Background and Theoretical Framework

Role of Interactive Discourse in Cooperative Groups

Educational researchers have emphasized the importance of student interactions in cognitive restructuring during group work. Group feedback and sharing of ideas help students in a group reformulate ideas and construct new knowledge that they might not have established on their own (Slavin, 1977). Wittrock (1974) focused on the active role that the learner plays when reformulating this information. Wittrock's model, which emphasizes the generative process of learning in which the learner is able to link new information to prior knowledge, has been applied by Webb (1980) in understanding learning in group settings. In these situations, students can help each other to evaluate their existing knowledge in light of new information and to alter or replace the existing knowledge if needed. More recently, in a review of literature relating to the role of discourse in group work, Nussbaum (2008) provides evidence that cognitive elaboration remains an important perspective for researchers seeking to understand collaborative discourse. Nussbaum concludes that collaborative discourse is most likely to lead to improvements in students' understanding of content when they have the opportunity for sustained practice with instructional norms that promote elaboration. In this context, the long history of cooperative learning as a "success story" (Slavin, 1996) can be seen as the classic case of a long-term intervention aligned with principles of cognitive elaboration.

Research has also demonstrated that successful cooperative learning is achieved by interactions and co-regulated engagement in the shared problem space (Roschelle & Teasley, 1995). Interactivity, which occurs when more than one student contributes to the discussion, has been identified as one of the necessary components for highly effective cooperative learning (Reusser, 2001). The level of interactivity among peers is dependent both on the frequency of interactions and also the extent to which these interactions influence the peers' cognitive processes. In our view, for the group process to produce changes in an individual's previously held ideas or beliefs, the individual must engage in argumentation with other group members to reformulate and restructure those ideas and opinions.

The risk that some students will not be engaged in the group process is always a concern for real-world implementations of cooperative learning, and researchers have suggested that multiple-ability grouping is preferable to ability-matched grouping, because low achievers often do not have much success in ability-matched groups (Evans, 1991; Slavin, 1990). However, a recent study of student discourse during a high school ecology project showed that knowledge was constructed in a meaningful and efficient way only between students with similar abilities (Rozenszayn & Ben-Zvi Assaraf, 2011). Esmonde's (2009) excellent review of mathematics education and cooperative learning introduces the idea of intersubjectivity and reinforces the proposition that fruitful research in this area will move away from a focus on group composition and toward an examination of the nature of students' participation. As will be shown below, argumentation has sufficient importance for science learning that an examination of student contributions to the co-construction of arguments within a group can provide useful insights.

The Importance of Argumentation in Science Learning

An argument can be thought of as the justification of claims with empirical evidence and reasoning. Argument construction can be either individual or social, and the two are often related in science learning. As Jiménez-Aleixandre and Erduran (2008) point out, "social dialogue offers a way to externalize internal thinking strategies embedded in argumentation" (p.12). Examination of the role of argumentation in science discourse has been gaining prominence over the past two decades. For example, some researchers have found that the teaching of argumentation strategies can improve the quality and quantity of student arguments at the elementary school (McNeill, 2011), high school (Osborne, Erduran, & Simon, 2004; Venville & Dawson, 2010), and college level (Nussbaum, Sinatra, & Poliquin, 2008; Yu & Yore, 2012). Others have investigated how the lack or presence of argumentation can have a corresponding impact on science learning. Duschl and Osborne (2002) argue that an absence of dialogical argumentation in the classroom could result in a reduction in science learning. Research has demonstrated the positive effect of argumentation on understanding of science concepts and improving reasoning skills in elementary school children (Mason, 1996; Mercer, Dawes, Wegerif, & Sams, 2004; Simon & Maloney, 2007), high school students (de Lima Tavares, Jiménez-Aleixandre, & Mortimer, 2010; Jiménez-Aleixandre, Rodrìguez, & Duschl, 2000; Zohar & Nemet, 2002), and college students (Aydeniz, Pabuccu, Cetin, & Kaya, 2012). These studies employed qualitative analysis of student utterances during argumentation or quantitative pre/post-test designs to explore students' knowledge gain. The results indicate that an increase in students' knowledge of science itself can be attributed to their engagement in argumentation.

Considering the importance of argumentation to science learning, research findings that demonstrate students' struggles with argumentation highlight the need to create opportunities for students to develop strong argumentation skills. For example, students have trouble explaining phenomena based on data (McNeill & Krajcik, 2007; Sandoval & Millwood, 2005) and they often do not provide reasoning to justify the claims that they do make (Kuhn & Reiser, 2005; McNeill & Krajcik, 2007). These studies have demonstrated that the nature of the task or structure of the class activities can have a substantial impact, either facilitating or hindering opportunities for students to engage in argumentation. One approach to enhancing student performance is to provide students with the opportunity to work in a group setting in which evaluation of scientific evidence and argumentation is employed. Improvements in argumentation skills can be achieved via such group discussion even without teacher intervention, as shown by research with South African high school students (Lubben, Sadeck, Scholtz, & Braund, 2009). In addition to supporting the growth of argumentation skills, allowing students to engage in argumentation collaboratively is thought to help students correct incorrect ideas by building consensus (Berland & Lee, 2012) and enhance students' scientific reasoning and understanding of scientific concepts (Osborne, 2010). Amigues (1988) showed that cooperative student groups performed better than individuals on science activities and Sampson and Clark (2009) reported that, even though groups did not produce better arguments than students who worked individually on initial tasks, students who had been part of collaborative groups during those initial tasks did better on mastery and transfer tasks than did students who had worked alone. These results are again consistent with the general observation that cooperative learning methods have the potential to enhance student learning (Barron, 2003; King, 1992, 1998; Springer et al., 1999).

Most of the studies on argumentation have been conducted in K-12 settings on biological science concepts (e.g., Berland & Hammer, 2012; Zohar & Nemet, 2002) and socio-scientific issues (e.g., Evagorou & Osborne, 2013; Venville & Dawson, 2010). Although argumentation studies in university-level chemistry are limited, recent work used argumentation as a lens to examine student discourse in physical chemistry classrooms. These studies analyzed students' conceptual progress and uncovered normative classroom practices (Cole et al., 2012), and also investigated how students develop particulate-level justifications for claims in thermodynamics (Becker et al., 2013). Another study explored the impact of an argumentation-based pedagogical intervention on general chemistry students' conceptual understanding of gases (Aydeniz et al., 2012). These studies do not explicitly examine the nature of student contributions and production of arguments within small group discourse. Our study investigates the nature of college-level general chemistry student discourse with respect to student participation in the construction and co-construction of arguments within the PLGI cooperative learning environment. Our focus is on un-mediated group discourse—that is, discourse that occurs when the group is working alone, unassisted by an instructor or other external facilitator. When working in a classroom with multiple small groups, un-mediated group discourse is frequently the predominant experience that students have as "group work." In addition, this approach enables us to directly investigate a concern that many instructors may have: when left on their own, groups may be likely to go "off track" and generate incorrect answers, or be dominated by one person who does all of the work and simply tells the other students "the answers." Our concerns regarding the nature of participation and our appreciation of the role of argumentation in promoting science learning lead directly to our research questions:

1. How frequently are various levels of individual and co-constructed argumentation observed within small student groups?
2. What patterns of participation in argumentation by individual students are observed in these groups?
3. To what extent do students in small groups resolve originally incorrect claims?

Collectively, these address our central question, "What are the characteristics of student group argumentation in the Peer-Led Guided Inquiry sessions of a General Chemistry I course, for arguments without peer leader intervention?"

Method

Analytical Framework: Toulmin's Argumentation Scheme

There are many analytical frameworks (Sampson & Clark, 2008) that can be used to assess the quality of student argumentation. For example, Enderle, Walker, Dorgan, and Sampson (2010) recently developed an observation protocol that focuses both on social interactions as well as argument structure to assess the quality of arguments in the classroom. One approach that has been widely used by science educators for the defining and examining of arguments is Toulmin's argumentation scheme, presented in his seminal work, The Uses of Argument (Toulmin, 1958). According to Toulmin's model (Figure 1), an argument contains several specific components. The *claim* is the conclusion at which one arrives. The *data* consist of evidence, information, facts, or procedures that lead to the claim that is being made. The *warrant* explains how the data or evidence leads to the claim. These three components (claim, data, warrant) are essential and constitute the *core* of the argument. Stronger arguments contain a *backing* that explains why the warrant has authority and provides the validity for the core of the argument. Additional components that may be present in more complex arguments are the *rebuttal* (a counter claim or a refutation of any of the components of the argument) and a *qualifier* (a limiting statement describing the conditions under which the claim holds true). In the student discourse analyzed for this work, these additional components were relatively uncommon, with qualifiers almost completely absent.

One criticism of Toulmin's argumentation scheme is the challenge in identifying and differentiating between warrants and backings (Keith & Beard, 2008). Others have critiqued it for the lack of warrants and backings in particular educational settings (Naylor, Keogh, & Downing, 2007). Toulmin's argumentation framework has also been criticized for catering better to a "monologue" rather than a "dialogue" (Plantin, 2005), for lacking the details needed to analyze dialectical arguments due to its general and wide categories (Duschl, 2008), and for losing the overall meaning when analyzing arguments (Furberg & Arnseth, 2009). Researchers also have found that the determination of whether a particular statement should be considered as a claim, data, warrant or backing is context dependent (Kelly, Druker, & Chen, 1998), which can be a potential drawback. Toulmin's argumentation scheme has also been criticized as catering to short

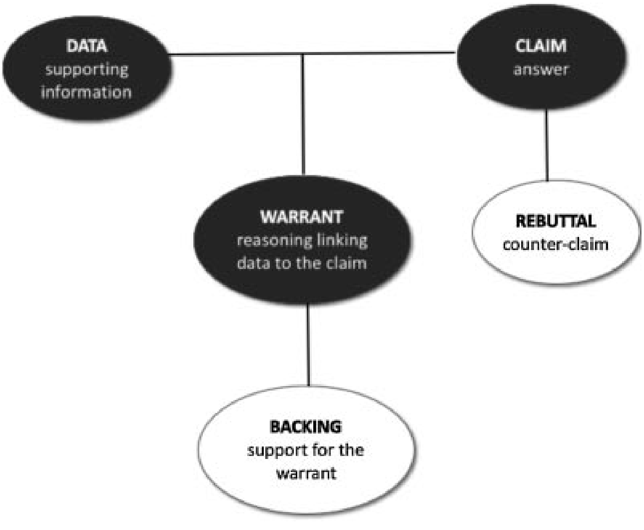

Figure 1. Toulmin's argumentation scheme.

argument structures (Chang & Chiu, 2008); however, this "flaw" is an ideal fit for the short (1–5 minutes) arguments occurring in the PLGI setting. A recent review on argumentation frameworks (Nussbaum, 2011) has also critiqued Toulmin's argumentation scheme as more suitable for determining the structure of arguments, and not the best tool for sophisticated arguments containing multiple schemes and/or moral reasoning.

Notwithstanding these criticisms, Toulmin's argumentation scheme has been successfully used to analyze argumentation in a broad spectrum of educational contexts, including studies in mathematics (Rasmussen & Stephan, 2008; Stephan & Rasmussen, 2002; Weber, Maher, Powell, & Lee, 2008), science (Becker et al., 2013; Cole et al., 2012; Foong & Daniel, 2011; Jiménez-Aleixandre et al., 2000; Sampson & Clark, 2011; Zohar & Nemet, 2002), and English (Mitchell, 1996). Toulmin's argumentation scheme has also been used as an analytical tool for the assessment of student work (Gotwals & Songer, 2010) and the quality of arguments (Abi-El-Mona & Abd-El-Khalick, 2011). Recently, Erduran et al. (2004) developed and applied Toulmin's argumentation scheme to the analysis of science discourse in middle school science classrooms. The analytical framework that they use to examine the strength of student argumentation is referred to as Toulmin's Argumentation Pattern (TAP). This approach has been modified by other researchers to better fit their specific research purpose. In a study designing and validating an instrument to assess younger students' arguments, Evagorou, Papanastasiou, and Osborne (2011) modified the TAP framework with alterations such as defining the lower levels as "contradicts the claim or no response (level 0)," and "appeal to authority (level 1)" and defining the highest level to be the "most convincing argument containing data, warrant, and rebuttal (level 4)." Evagorou, Sadler, and Tal (2011) have also discussed the implications of modifying TAP to incorporate the number of pieces of evidence presented in arguments as an additional measure of the argument quality.

For our purposes, it is the presence of the components of argumentation and who is contributing these components that is of interest. Thus, we extend and complement these previous efforts by applying our own modification of TAP to analyze argumentation in small groups of students in a post-secondary general chemistry context. Toulmin's argument components are exactly relevant to evaluate the presence or absence of data and reasoning in students' discourse. Toulmin's scheme facilitates the identification of which component was provided by which student, revealing participation patterns. In comparison to alternative frameworks such as Walton's Dialogue Theory, Toulmin's argumentation scheme is a better choice for our current research focus (Nussbaum, 2011).

Coding Scheme Based on Toulmin's Model

Previous investigators have noted that developing a coding scheme based on Toulmin's argumentation model can be a challenging task (Erduran et al., 2004). As mentioned previously, the process of identifying the various components of an argument's core and backing can be difficult and context dependent; a robust coding scheme is therefore essential. We developed the coding scheme used in this work from definitions provided by Toulmin for each component and from examples of coded arguments in previous studies (Erduran et al., 2004; Stephan & Rasmussen, 2002). The goal of this initial coding was to identify the components of each argument (data, claim, warrant, backing, rebuttal) and to note which member of the group provided the component. This part of the analysis is based on the approach first described by Erduran et al. (2004) and used successfully in undergraduate chemistry courses by others (Becker et al., 2013; Cole et al., 2012); it is also similar to the approach used recently to analyze argumentation in a socioscientific context by middle school pairs (Evagorou & Osborne, 2013). As described in these previous efforts, we looked for cues in the student utterances such as "so" or "because" to assist in determining how a particular contribution was operating within a specific exchange.

Before describing the coding scheme in more detail, we first present a description of the written materials used as the basis and catalyst for group discourse in the PLGI setting. These paper-and-pencil materials, referred to as ChemActivities (Moog & Farrell, 2008), are designed to be used by students in groups of three or four, with the instructor (or peer leader) serving as a facilitator who listens to the discussion and intervenes only when necessary. The activities are structured to guide students through an investigation of presented data, figures, or verbal descriptions to build chemical concepts. Many of the guiding questions contain scaffolds such as "why" or "explain" that elicit explanations for the answers. For example, the activity dealing with atomic and ionic radii begins with a reminder about the previously established periodic trends in first ionization energy. Then a data table indicating the valence shell, core charge, and atomic radius of numerous atoms is presented. Through a series of guiding questions, the students are led to recognize the trends in atomic radii across a period and down a column, and they are prompted to articulate explanations for these trends in terms of the atom's core charge and valence shell. The students are then asked to apply these concepts by predicting the radii of three atoms not listed in the table, and explaining how they arrived at their estimates. Thus, these materials are not typical drill worksheets or collections of typical homework or exam questions; rather, they are specifically structured to promote analysis and interpretation, discussion, and student articulation of reasoning.

We used the following rubric to anchor the coding procedure. An answer offered by a student to a ChemActivity question, whether simple (e.g., two electrons) or more complex (e.g., ionization energy increases as the effective nuclear charge increases) is considered a *claim*. The chemical information the student used as evidence to arrive at that claim is labeled as *data*. For instance, a balanced chemical equation, a mathematical formula, or a variable from a mathematical formula that the students used as evidence to support a claim would be identified as data. In some cases, a claim or the data may be presented as part of the question in the activity. A *warrant* is an explanation of how the data or evidence leads to the stated claim; scientific reasoning and explanations offered by students using associated course concepts to support their answers are considered warrants. Sometimes warrants are mathematical, for example, a mathematical operation including an explanation of the operation. An elaboration of an explanation, an offer of valid common patterns, or explicit reference to chemistry theories and laws or previously learned concepts by the students in order to expand their warrants is labeled as a *backing*. Finally, a counter claim offered by one student to oppose a claim offered by another student is considered a *rebuttal*. Within an argument, components other than the claim can also be rebutted, and those are also labeled as rebuttals (Stephan & Rasmussen, 2002; Toulmin, 1958). As mentioned previously, *qualifiers* were extremely rare in this study, and were not coded.

The presentation of a claim on a new topic or in response to a question in the ChemActivity defined an *episode*. Episodes and arguments are not synonymous; in order for an episode to be classified as an argument, it must contain all of the elements of the core of an argument. Thus, only those episodes of student discussion that contain at least a claim, data, and a warrant are considered to be arguments. Although these three core components were not always articulated in this order, all were required to be present for the episode to be classified as an argument. In the higher levels of argumentation, a backing and/or rebuttal would also be present. Note that this approach differs from the TAP framework originally described by Erduran et al. (2004) in which an argument need not contain all three of the core components that we require. This is a reflection of the different focus of our work on the production of full arguments rather than "the quality of opposition or rebuttals in the student discussions" (Erduran et al., 2004). In addition, in our study, almost all of the episodes (and arguments) were directly prompted by the questions in the ChemActivities.

The following excerpt of an episode that constitutes an argument provides an illustrative example of the application of this coding scheme. Codes are shown in parenthesis in capital bold face letters (**CODE**). The small group (Group A) consists of four students: Scott, Joe, Mike, and Ron. All names used are pseudonyms. In this episode, the students are working on a ChemActivity concerning dipole moments. They are answering the question (number four): "Why is the dipole moment zero for CCl_4?" The appropriate Lewis structure for CCl_4 had been drawn by the students previously on the activity page. The equation ($\mu = q \times d$) for calculating the dipole moment had also been introduced earlier in this activity.

Dipole Moment Argument

[00:36:26.10] Joe: Yeah, what number are we on?
[00:36:24.25] Scott: On 4.
[00:36:26.17] Joe: Four.
[00:36:33.28] Mike: Oh yes
[00:37:04.05] Mike: So it's zero **(CLAIM)** because of distance, right? **(DATA)**
[00:37:06.20] Scott: Because of distance between the center of charge is zero. Yeah.**(WARRANT)**
[00:37:08.26] Joe: What? Why is it?
[00:37:19.12] Mike: Just like the other . . . the CO_2 because there's no distance between the center of the charges. **(BACKING)**

Although these students may not use the language that experts would employ or provide full explanations of their thinking, they do articulate the various components of an argument. Mike begins by restating the *claim* presented in the question that the dipole moment of CCl_4 is zero. (In general, the claim is not included as part of the question or prompt in the ChemActivities, although it is occasionally present, as in this example.) In addition to stating the claim, he provides the evidence for this claim: "the distance." Therefore, "because of distance" is coded as *data*. Scott then expands on that evidence of distance by explaining that the claim is true because the distance between the centers of charge is zero. Since Scott is explaining how the evidence led to the claim, this statement is labeled as a *warrant*. At the very end, Mike validates the core of the argument by illustrating that their claim and the reasoning is justified: the CO_2 molecule they had previously analyzed also had a distance between the centers of charges of zero, and they had previously concluded that in that case the dipole moment in CO_2 is zero. Here he is providing a *backing*—validating the argument by referring to a previous example in which analogous reasoning holds true.

An Analytical Framework for Characterizing Argumentation in Small Groups

The application of the coding scheme as demonstrated above provides the basis for characterizing the various arguments that are produced by the student groups. The two factors that are of particular relevance in our investigation are (1) whether any additional components (backing, rebuttal) beyond the core are present, creating a "stronger" argument; and (2) whether more than one student contributes to the construction of the argument. As mentioned previously, earlier work by Erduran et al. (2004) on argumentation in science discourse played an important role in the development of an appropriate analytical framework for characterizing argumentation in our context. The various levels of argumentation that Erduran and coworkers identified in TAP served as a starting point for our framework, presented in Table 1. However, as with some previous workers (Evagorou et al., 2011a, 2011b), several adaptations were necessary to deal with the particular circumstances and goals of the study. Our focus is on the production of arguments that

Table 1
Framework used for assessing the quality of argumentation

Condition	Level	Description
Individual Arguments	Level I1	Claim, data, warrant(s) provided by one student
	Level I2	Claim, data, warrant(s) provided by one student, backing(s) provided by the same student
Co-constructed Arguments	Level C1	Claim, data, warrant(s) provided by more than one student
	Level C2	Claim, data, warrant(s) and backing(s) provided by more than one student
	Level C3	Claim, data, warrant, and a rebuttal provided by more than one student (with or without backing)

contain, minimally, all three of the core components, whereas Erduran et al. determined the level of argumentation primarily by the strength and presence of rebuttals, independent of the presence of all of the core components. For example, Erduran et al.'s Level 1 consists of an argument containing only a claim, which we identify as an "episode" but not an "argument." Thus, an episode that included only a claim by itself was not assigned a Level, and was not counted as an argument. In addition, an important characteristic of an argument in our context is whether it is provided completely by one student (an individual argument) or involves more than one member of the group (a co-constructed argument). Thus, we differentiate between these two types of argument in our scheme, with Levels I1 and C1 designating arguments that include the core only, and are either individually presented (I1) or co-constructed (C1). Erduran et al.'s Level 2 is somewhat analogous to our Levels I2 and C2: backings must be present. The presence of a backing indicates a stronger argument, because a backing explains why the warrant has authority and may also elaborate on the reasoning used to arrive at the claim from the data. However, Erduran et al. did not require all three components of the core to be present in Level 2 as we do for all of our Levels. Because the contrast between individual and co-constructed arguments was not a focus of study for Erduran et al., the distinctions between their Levels 3, 4, and 5 were based on the strength of the rebuttal presented in an argument. This approach is apt in their context because their study focused on debatable socioscientific issues, with many different perspectives present and more than one valid answer possible. Our study, however, involves chemistry questions and problems that generally anticipate one correct answer, although there may be multiple ways to arrive at that answer. Thus, there were relatively few rebuttals provided in our context, and categorizing the arguments based on the strength of the rebuttals was less relevant than simply considering whether or not a rebuttal was present. Therefore, Level C3 describes a co-constructed argument that contains a rebuttal from a member of the group other than the individual who presented the rebutted claim. Level I3 for individual arguments is not included in our coding scheme because there were almost no instances of self-rebuttal.

Note that the argument presented above (Dipole Moment Argument) is classified as Level C2 because it contains a backing and is a co-constructed argument, with more than one student contributing at least one component.

Sample

This investigation focused on two groups of four students facilitated by two different peer leaders in separate rooms. None of the students were chemistry majors, but all were taking the course to satisfy a major requirement. Group A consisted of four white male students (Scott, Mike, Joe, and Ron) ranging from sophomore to senior in class year, with an undergraduate peer leader majoring in biochemical science and peer leading for the first time. Group B consisted of three female students and one male student, ranging from freshman to sophomore in class year. Two of the female students are Asian (Michiko and Janet) and the other female student (Monifa) is Black. The male student (Sam) is also Black. The peer leader of Group B was a chemistry graduate student with some prior peer leading experience. Maximum diversity sampling (Patton, 2002) was used to select the two focal groups. These groups were chosen because of their difference in diversity with respect to sex, race/ethnicity, and class year. Both groups were mixed ability based on SAT scores; however, the coders did not have access to the SAT scores or the student final grades in the course at the time of coding the discourse and classifying the arguments. Each group's student composition remained constant throughout the semester. Each student was assigned a role within the group each week. All the students except for Sam and Janet in Group B were present for all 12 peer-led sessions. Sam was absent for four sessions and Janet was absent for one session.

The peer leaders attended a training course that met throughout the semester under study. The instructor for this one-credit course was a chemistry faculty member with substantial experience and expertise in small group facilitation. During the first hour of each 2-hour weekly session, all of the peer leaders for the general chemistry course worked in small groups on the ChemActivity that the students would encounter in the upcoming peer-led sessions. The second hour was devoted to discussion of potential student difficulties and misconceptions and possible strategies to deal with these issues.

Data Collection

The 50-minute PLGI sessions each included about 20 students and were held on Friday of each week. The two focal groups (A and B) were videotaped during each of their 12 weekly peer-led sessions during the Spring 2008 semester. The whole classroom was videotaped in order to capture the dynamics of the class and also to minimize any imposition on the focal group. The audio portion of each session was transcribed for analysis. Informed consent was obtained from all students and the two peer leaders prior to videotaping the sessions.

Each session began with a quiz and a brief procedural introduction by the peer leader indicating which sections of the designated ChemActivity were to be completed. The students then worked on the activities in their small groups. During a typical session, the peer leader would facilitate one or two whole-class discussions of important or difficult questions or concepts, and the session would end with a written group report that included some reflection on the group's performance during the session. The time spent in small group work each week ranged from 20 to 35 minutes. There was only a difference of a few minutes between Groups A and B for any given week; this small difference reflected the varying amount of time the respective peer leader spent on the quiz, introduction, and whole class discussions. Thus, both groups under study had roughly equal amounts of time available for small group work. The number of arguments that occurred during this small group work time was tallied as the frequency of arguments for our study.

Coding and Data Analysis

The argumentation data for each session were obtained by reviewing both the video recording and the corresponding transcript. Episodes were initially identified on the transcripts while watching the videos, and the corresponding student statements were coded as described in the coding scheme above. Finally, the presence (or absence) of an argument was established for each episode, and a Level was assigned based on the framework presented in Table 1. A second rater, blind to the coding of the initial rater, analyzed two transcripts containing about 10% of the total number of arguments. Three different aspects of the analysis were compared to check inter-rater reliability. First, the identification of the presence of an argument was examined; there was 100% agreement between the two raters. Next, the question of whether an argument was co-constructed or individual was investigated; the percentage of agreement was 94%. Finally, on the assignment of a Level to the arguments, there was 83% agreement.

The arguments within the groups were also classified to indicate whether or not the peer leader was involved in the process. Roughly 80% of the arguments observed within these two groups were produced without peer leader intervention. However, the analysis presented in this study involves *only* those episodes in which students worked on their own, without any interaction with the peer leader during the discussion. That is, all of the results presented here are for episodes and arguments in the absence of peer leader involvement or intervention. Our study of arguments with intervention is presented elsewhere (Kulatunga & Lewis, 2013).

Results and Discussion

Before we address our research questions, we first examine all of the episodes (including those that are not arguments) to determine the extent to which students support the claims that they put forward with at least some data. This is an important issue to address because some previous reports have shown that students often offer claims without any data as support (McNeill & Krajcik, 2007; Sandoval & Millwood, 2005). Therefore, it is important to analyze all of the episodes to get a measure of how often the students offer support for their claims—even if no warrant is present. For this purpose, all of the episodes in the two focal groups were tallied and placed into three categories: claims, claims with data only, and arguments (Table 2).

As shown in Table 2, only a small percentage of claims are presented in this setting without any data to support them (Group A: 11%, Group B: 14%), in contrast to the previous work mentioned above. Notably, the majority of claims are supported by both data and warrants; that is,

Table 2
Distribution of student discourse episodes

	Group A				Group B			
ChemActivity	Claims Only	Claims and Data	Arguments	Total	Claims Only	Claims and Data	Arguments	Total
Balancing Chemical Equations	3	1	5	9	5	2	8	15
Limiting Reagent	3	1	8	12	2	1	10	13
Coulombic Potential Energy I/Enthalpy of Atom Combination	1	1	14	16	3	1	9	13
Specific Heat	0	2	11	13	0	0	9	9
Coulombic Potential Energy I/Shell Model I	0	1	8	9	0	3	4	7
Shell Model II	0	0	9	9	1	5	10	16
Atomic Size	0	1	11	12	2	4	11	17
The Ionic Bond	0	2	11	13	2	4	9	15
Lewis Structures I and II	4	3	9	16	1	4	9	14
Lewis Structures III and IV	2	4	6	12	0	4	6	10
Dipole Moment	0	0	8	8	2	4	7	13
Intermolecular Forces	2	3	5	10	3	1	4	8
Total	15 (11%)	19 (14%)	105 (75%)	139	21 (14%)	33 (22%)	96 (64%)	150

most of the episodes resulted in the construction of arguments (Group A: 75%, Group B: 64%; overall: 69%). These results are promising because they indicate that the students in this study are generally providing evidence and reasoning to justify their claims as part of their discourse, and they are doing so in the absence of direct prompting or intervention from the facilitator. Previous work (Jiménez-Aleixandre et al., 2000) suggests that students understand and learn chemistry concepts better when they are able to support their answers with evidence and reasoning instead of just making unsupported claims. From this perspective, this PLGI setting appears to provide a productive learning environment. We now proceed to a discussion of each of the research questions.

How Frequently Are Various Levels of Individual and Co-Constructed Argumentation Observed Within Small Student Groups?

Tables 3 and 4 provide, for Groups A and B respectively, the number of arguments at each Level without peer leader intervention for each PLGI session. The tables also show the total number of arguments at each Level for each of the groups; this distribution is also provided as a bar graph for comparison purposes in Figure 2. Although there is a statistically significant difference in the proportion of arguments that are co-constructed between the two groups ($n = 201$; $d = 1$; $\chi^2 = 4.5$; $p < 0.05$), both groups have a substantial number of arguments that are produced by individuals and that are co-constructed, with a large majority of the arguments containing only the core.

These data indicate that a significant majority of the arguments without peer leader intervention were co-constructed in both groups. This result suggests that, in most cases, when one student provides a claim as part of an argument, at least one other member of the group contributes a component of the argument—data, warrant, backing, or rebuttal. In fact, of the 289 total episodes produced by these two groups throughout the semester, 145 (50%) resulted in co-constructed arguments, with the other 50% of episodes being either individual arguments (56 or 19%) or non-

Table 3
Distribution of the levels of argumentation for Group A without peer leader intervention

	Level of Argumentation					
ChemActivity	Level I1	Level I2	Level C1	Level C2	Level C3	Total
Balancing Chemical Equations	2	—	2	1	—	5
Limiting Reagent	4	—	3	1	—	8
Columbic Potential Energy I/ Enthalpy of Atom Combination	3	—	6	2	3	14
Specific Heat	6	1	3	—	1	11
Columbic Potential Energy I/ Shell Model I	4	—	2	—	2	8
Shell Model II	1	1	7	—	—	9
Atomic Size	2	1	7	—	1	11
The Ionic Bond	1	1	5	—	4	11
Lewis Structures I and II	2	—	6	1	—	9
Lewis Structures III and IV	3	—	2	—	1	6
Dipole Moment	3	—	2	3	—	8
Intermolecular Forces	1	—	2	—	2	5
Total	32 (30%)	4 (4%)	47 (45%)	8 (8%)	14[*] (13%)	105
Total by Argument Type	Individual 36 (34%)		Co-Constructed 69 (66%)			

[*] Four out of the 14 arguments in Level C3 also contained backings.

Table 4
Distribution of the levels of argumentation for Group B without peer leader intervention

	Level of Argumentation					
ChemActivity	Level I1	Level I2	Level C1	Level C2	Level C3	Total
Balancing Chemical Equations	—	—	7	—	1	8
Limiting Reagent	2	—	5	1	2	10
Columbic Potential Energy I/ Enthalpy of Atom Combination	3	—	2	1	3	9
Specific Heat	1	—	5	—	3	9
Columbic Potential Energy I/ Shell Model I	—	—	3	1	—	4
Shell Model II	2	—	4	1	3	10
Atomic Size	1	—	5	—	5	11
The Ionic Bond	3	—	3	1	2	9
Lewis Structures I and II	2	—	6	—	1	9
Lewis Structures III and IV	2	—	3	1	—	6
Dipole Moment	1	—	3	1	2	7
Intermolecular Forces	3	—	—	1	—	4
Total	20 (21%)	0	46 (48%)	8 (8%)	22* (23%)	96
Total by Argument Type	Individual 20 (21%)		Co-constructed 76 (79%)			

*Five out of the 22 arguments in Level C3 also contained backings.

argument episodes (88, or 30%). Thus, the typical discourse in these groups can be reasonably characterized as discussion, rather than a monologue from one individual. This is of significance because some instructors (and students) may be concerned that a single student in the group (generally, the "strongest" student) would generate all of the answers and the other group members would simply write them down. These data suggest that is not the dominant paradigm for working

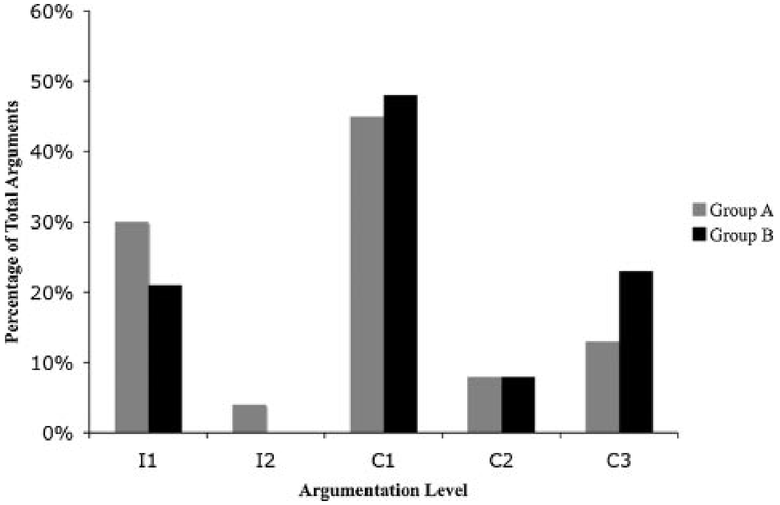

Figure 2. Distributions of arguments by Level for Groups A and B.

through the ChemActivities in this setting. We further address the issue of who is contributing to the arguments and in what ways in our examination of research question #2 below.

The distributions shown in Figure 2 indicate two additional points of interest. First, we see that arguments at Levels I1 and C1, which contain only the core, are more frequent than the stronger arguments that contain backings and/or rebuttals (Levels I2, C2, C3) for both groups. For Group A, 79 out of 105 arguments (75%) and for Group B, 66 out of 96 arguments (69%) contain only the core. The relative number of arguments with backings is very low: 16/105 for Group A and 13/96 for Group B. These low percentages of arguments with backings (15% for Group A; 14% for Group B) indicate that students do not often elaborate on their reasoning, validate their explanations by providing relevant chemistry theories or laws, or explicitly articulate that they are applying previously learned concepts to new examples. However, a higher percentage of co-constructed arguments (24/145 = 17%) contain backings than do individual arguments (4/56 = 7%). This finding suggests that students are more likely to elaborate on their reasoning when co-constructing arguments in a group rather than making individual arguments. Thus, promoting the use of collective argumentation in student groups may be one strategy to increase the presentation of backings and thereby strengthen student arguments.

The low percentage of Level C3 arguments indicates that these students did not frequently provide rebuttals during group discussions. Other research also has discovered a lack of rebuttals in student arguments (Chen, Lin, Hsu, & Lee, 2011). Perhaps not surprisingly, rebuttals generally were observed when incorrect claims were put forward. However, relatively few incorrect claims were presented, resulting in the low number of rebuttals observed. It is important to note, however, that in almost all cases in which incorrect claims were presented, the group was able to resolve them and agree on the correct claim through their interactions (including the presentation of a rebuttal). This finding will be discussed further below when the resolution of wrong claims is addressed.

What Patterns of Participation in Argumentation by Individual Students Are Observed in these Groups?

There are two contexts in which to examine this question: for individual arguments and for those that are co-constructed. Because the patterns of involvement within a group may be different in these two distinct circumstances, we examine each context separately, and then compare the results. Tables 5 and 6 present the data concerning individual arguments (Levels I1 and I2) within each group.

One interesting result is that there is no necessary correlation between the production of individual arguments and the overall performance of the student in the course. That is, chemistry content knowledge as measured by the final course grade is not necessarily the primary indicator of who is presenting individual arguments. Note that in Group A (Table 5), the final course grade tends to mirror the relative contribution of individual arguments; however, this is not the case for

Table 5
Distribution of individual arguments (n = 36) among Group A members

Student	Number of Sessions	Individual Arguments Offered	Percentage of All Individual Arguments	Average Arguments per Session	Final Course Grade
Scott	12	22	61	1.8	A−
Mike	12	8	22	0.67	B
Joe	12	4	11	0.33	B
Ron	12	2	6	0.17	C

Table 6
Distribution of individual arguments ($n = 20$) among Group B members

Student	Number of Sessions	Individual Arguments Offered	Percentage of All Individual Arguments	Average Arguments per Session	Final Course Grade
Janet	11	8	40	0.72	C−
Michiko	12	8	40	0.67	B−
Sam	8	2	10	0.25	F
Monifa	12	2	10	0.17	A−

students in group B (Table 6). There are two important and related points to be made with respect to this observation. First, the relative "strength" of a student does not determine the relative contribution of "answers" from that student that include evidence and reasoning. In Group A, the individual (Scott) who contributes 61% of the arguments has the highest chemistry content knowledge as measured by the final course grade. In Group B, the individual with the highest final course grade (Monifa) contributed the fewest individual arguments per session. Second, the two groups did not behave similarly. Although not a focus of this study, the interplay between group composition, individual personality, background knowledge and other factors likely influences the patterns of contribution of individual arguments. Note that Group A is a relatively homogeneous group of four white males and Group B is a heterogeneous group with respect to race and gender, with Sam, who did not pass the course, as the only male student among three females. Previous research (Webb, 1984) has shown that group compositions that isolate one gender can lead to interactions that can be detrimental to learning. Thus, demographic isolation could be a factor that hindered Sam's participation in argumentation and his success in the course.

Examining the extent to which the particular demographics of these two groups are responsible for the differences in contributions of individual arguments is beyond the scope of this work. However, one final important point can be made. Our framework, based on Toulmin's argumentation scheme and a modification of TAP targeted to our research questions, enables us to identify and codify the differences between these two groups.

We now turn our attention to an examination of the participation patterns in co-constructed arguments. Figures 3 and 4 present the relative contributions of each student for each argument component as a percentage of the total number of occurrences of that component in all of the arguments without peer leader intervention. Several differences are apparent from a comparison of these two figures. The pattern of contributions in Group A is essentially the same for all of the components (with the exception of a variation for rebuttals, which will be addressed below) whereas in Group B the pattern of contributions varies across the components. In Group A, Scott contributes the most for all of the components, with Mike second for all but the rebuttals. Joe and Ron each make about 10–15% of the contributions in all categories, with the exception of rebuttals where Mike rarely contributes and Ron provides 25% of the remarks. This pattern generally mimics the relative contributions of individual arguments described above. In contrast, the results for Group B are quite different. There is no constant pattern of contributions for the various components as there is for Group A (although Sam does generally contribute the least of any member of the group). The two female students, Janet and Michiko, who equally dominated the production of individual arguments also contribute substantially to each of the components of the co-constructed arguments, but Monifa also participates in the co-constructed arguments to a significant extent, much more than the 10% of individual arguments that she produced. Thus, in Group B the co-constructed arguments reflect a much greater degree of interaction and discourse

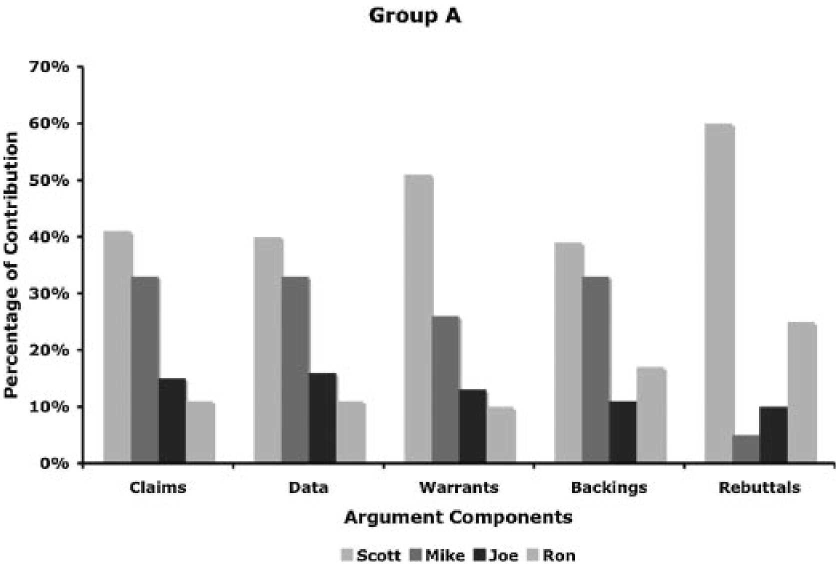

Figure 3. Individual contributions to co-constructed arguments in Group A.

involving all members of the group, whereas in Group A the co-constructed arguments tended to involve primarily two group members.

As noted above, one exception to the pattern of discourse in Group A involves the relative contribution of rebuttals. Scott provided about 60% of the rebuttals. This is consistent with his domination of the production of individual arguments and his high overall course grade. Scott's role as the dominant participant in arguments was due to the lack of contribution by other group

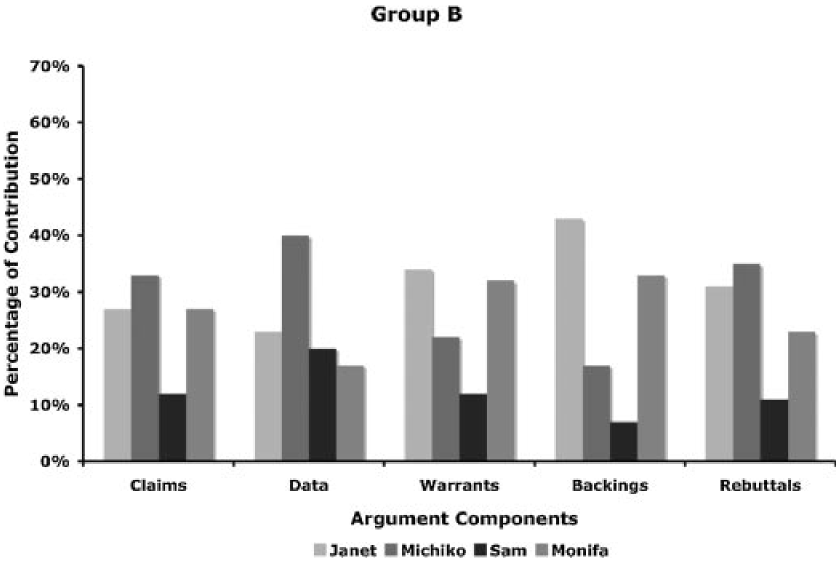

Figure 4. Individual contributions to co-constructed arguments in Group B.

members and was not reflective of an "alienating leader" who was not interested in what the other group members had to say or who imposed ideas on the others (Richmond & Striley, 1996). Scott had confidence in his understanding of the material, tended to produce answers and explain his thinking to the other males in his group, and when someone else provided a claim that he thought was incorrect, he would indicate his disagreement and provide his alternative claim. Of note, however, is that Mike is providing almost no rebuttals, but Ron is providing many more. In fact, Ron provides a greater percentage of rebuttals than any other argument component. Of interest is a comparison of Ron's pattern of contributions to those of Sam in Group B. Each of them provided only two individual arguments over the entire semester. They both provided a low percentage (about 10%) of contributions to the cores of co-constructed arguments—with the exception of Sam's contribution of about 20% of Group B's data. However, there is a difference in their relative contributions of backings and rebuttals, the elements of stronger arguments. Sam's contributions in these areas remain very low—the lowest contributions in Group B. But Ron has higher contributions in these areas than in the core components, and he replaces Mike as the second highest contributor of rebuttals in the group. Ron provided roughly equal numbers of rebuttals to claims made initially by Scott and by Mike. Of interest is the fact that Ron and Sam each had the lowest math SAT score in their respective groups; in the end, however, Ron earned a C in the course, and Sam did not pass. Ron's ability to contribute backings and rebuttals may be an indication of (or reflection of) his better grasp of some of the important concepts in the course. The relationship of the production of backings and rebuttals to overall success and understanding is an area for possible further investigation.

Both groups showed a discrepancy among group members with regard to the individual arguments; however, there are some differences in the participation patterns of the two groups. In Group A, one student (Scott) dominated in providing individual arguments in comparison to the other three group members. In Group B, two students (Michiko and Janet) took the lead in offering individual arguments and the other two were less engaged in individual arguments. Observations and coding of the videos showed how this difference in the participation pattern in the two groups resulted in Group B having more opportunity for collaborative argumentation. This discrepancy in the participation structure of the two groups also would explain the much higher percentage of Level C3 arguments with rebuttals in group B in comparison with group A (Figure 2). The Group B environment, where most of the members contributed some to the construction of arguments, unlike in group A where one member dominated most of the contribution, allowed a favorable atmosphere that promoted rebuttals. In Group B students also occasionally split into two pairs, with each pair working separately on a problem instead of as a whole group. This arrangement frequently promoted rebuttals between the pairs. This type of pairing of students within a group in cooperative learning has been observed previously (Daubenmire & Bunce, 2008). Our analytical framework enabled us to uncover these two different participation patterns. The ability to identify such patterns is important because it provides a context for implementation of an intervention that targets a specific participation pattern within a student group.

To What Extent Can Students in Small Groups Resolve Initially Incorrect Claims Without Peer Leader Intervention?

Although working in groups can have numerous benefits, one possible concern is that the students will generate incorrect answers or ideas that will create or reinforce misunderstandings and misconceptions. This may be of particular concern in situations such as ours, in which the instructor is not present and the peer leader is relatively inexperienced in both classroom management and in helping students develop content expertise.

To investigate the issue of resolution of incorrect claims, the original claim for each episode without peer leader intervention was identified as being correct or incorrect. For the episodes initiated by an incorrect claim, careful analysis of the transcript and video provided a basis for determining whether or not the incorrect claim was resolved by the end of the episode. Phrases such as, "I understand now," "I agree," "that makes sense," given by students after a lengthy discussion that included reasoning with a presentation of the correct claim were also used as indicators for resolution of wrong claims through argumentation.

As described previously, in some cases the ChemActivity contained prompts for explanations that supported the resolution of initially incorrect claims. In most cases, however, the students independently addressed the incorrect claim. When one student presented an incorrect claim, one or more other group members offered a rebuttal and provided reasoning as support for their rebuttal. Through this type of interaction, the group was able to not only arrive at a correct resolution but also helped each other better understand the underlying chemistry concepts. An example of this process is demonstrated in the following excerpt from a Level C3 argument from Group A, dealing with part of the "Atomic Size" ChemActivity described earlier. Here the students are working on the question, "Predict which is larger: the O^{2-} ion or the F^{-} ion?" given in the activity.

Ionic Size Argument

[00:31:17.26] Joe: Which is bigger? They're equal. Right? **(CLAIM-incorrect)**
[00:31:26.07] Scott: No. **(REBUTTAL)**
[00:31:27.10] Joe: Yeah. Eight and 2 would be 10 and then F is 9 and 1. **(DATA)**
[00:31:31.27] Scott: Um . . . Oxygen would be larger. **(REBUTTAL)**
[00:31:36.12] Joe: No.
[00:31:35.23] Scott: They can have the same, like, electrons but um. . .fluorine has more protons **(DATA)** so it's going hold its electrons closer. **(WARRANT)**
- - - - -
[00:32:02.00] Mike: They can't be the same as equal.
[00:32:06.17] Joe: Oh. So then . . .
[00:32:11.00] Mike: Let's find the core charge.
[00:32:15.19] Scott: ________. [inaudible]
[00:32:17.18] Joe: Isn't that what I'm saying? It's the same thing, right?
[00:32:19.24] Scott: Um . . .
[00:32:19.07] Joe: So they would both have 10.
[00:32:20.18] Scott: So the core charge for fluorine is going to be greater because it has more protons. **(BACKING. . .)** So you'll have oxygen only has 8. . .
[00:32:26.15] Joe: How do you know it has more protons?
[00:32:29.19] Scott: Because see 8 and 9, that's the number of protons it has.**(DATA)**
[00:32:30.29] Joe: Okay.
[00:32:31.25] Scott: Since this has 9 that means it's going to be a greater charge in the center so they're going to have . . . pretty much they're going to have the same number of electrons. **(DATA)** But since its core charge is greater it's going to pull 'em in closer so it's going to make it smaller. **(. . .BACKING)**
[00:32:46.04] Joe: Okay.

Joe begins this argument by making the incorrect claim that the sizes of the oxide and fluoride ions are the same. After Scott disagrees by saying "No," Joe provides some data for his claim by correctly calculating that both ions are 10 (referring to the total number of electrons). Scott rebuts Joe's claim at 31:27 and then provides both data (same number of electrons, different number of protons) and warrant (since fluorine has more protons, it would make the radius of the fluoride ion smaller) at 31:35. Mike suggests that the group consider the core charge as a possible way of approaching the issue. Joe does not appear to understand the concept of core charge, as is evident from his statements, "Isn't that what I'm saying? It's the same thing, right?" (32:17) and "So they would both have 10" (32:19). Scott helps Joe understand the difference between the core charge and the total number of electrons by again referring to the concept of protons at 32:20. This reference is the beginning of the backing, which Scott continues at 32:31, after Joe asks him how he knows that fluorine has more protons than oxygen. Scott explains the meaning of the atomic number, and can be seen pointing at the atomic numbers of oxygen and fluorine on the periodic table when he says, "Because see 8 and 9, that is the number of protons it has." This statement is labeled as data since it is some of the evidence that was used for the explanation (warrant) Scott gave above. At the end of the argumentation, (32:31) Scott includes the data that Joe originally provided—that oxide and fluoride ions have the same number of electrons—and then concludes by continuing the backing related to core charge that he began earlier. This statement validates the reasoning ("pull them in closer") for his rebuttal that oxide is the larger ion; therefore, it is labeled as a backing. This example demonstrates how the students resolved an incorrect claim through argumentation and also helped each other understand an important and challenging chemistry concept.

Examining the accuracy of all the claims provided by students during the 105 arguments without peer leader intervention in Group A, 22 were initially inaccurate. The students were able to resolve 20 out of these 22 inaccurate claims (91%) by engaging in argumentation without peer leader intervention. For Group B, 23 initially inaccurate claims were made during a total of 96 arguments. The students were able to resolve 20 out of these 23 inaccurate claims (87%) by engaging in argumentation without peer leader intervention. Overall, out of 201 arguments without peer leader intervention produced by these two groups, 196 of them (97.5%) ended with a correct claim being accepted by the group. Not only are relatively few incorrect claims generated, but the vast majority that do arise are resolved within the groups without intervention. These outcomes demonstrate the favorable impact of argumentation and small group interactions on students' independence in this cooperative learning environment.

Conclusions and Implications

We have presented a modification of the TAP framework for analyzing student discourse involving argumentation in small group settings based on both the strength of the argument and the extent of participation. This framework is a useful tool because it enables observation of group processes and individual participation patterns within cooperative groups, and thus can be applied in any context or discipline in which small group discussion is involved.

There are some limitations of the findings from our study. First, our findings are based on the argumentation analysis of only two groups of students. Our current and future work involves analyzing more groups in different peer-led sessions from different academic years. In addition, our analysis is limited to one type of cooperative class setting led by student peer-leaders, and is based on discourse concerning the specific chemistry topics covered in the peer-led sessions. Therefore, it is not certain that these findings can be generalized to any type of cooperative educational setting with different types of instructors, students, or content.

Previous researchers have warned that the benefits of group work are not universal (Bianchini, 1997, 1999; Cohen, 1984; Veenman, Denessen, van den Akker, & van der Rijt, 2005). Our analytical framework is a useful device for considering student participation patterns in a small group-learning environment. Our results reveal that students are not contributing equally in the construction of arguments, and that interactions among group members varied between the two groups. In one case, one of the students dominated in the construction of both individual and co-constructed arguments, while in the other instance two students participated more equally as leaders of individual arguments, and a third joined in substantially contributing to collective argumentation. The large difference in the percentage of arguments with rebuttals between the two groups reflects differences in student participation patterns.

Most of the argumentation that occurs in these small groups was co-constructed, suggesting that students are working cooperatively in this group setting. In addition, students provided a greater percentage of higher-level arguments with validations when co-constructing arguments rather than when doing so individually. Our findings add to the literature that has examined the impact of collaboration on constructing arguments such as a study by Sampson and Clark (2009) who found that even though groups did not produce better arguments than students who worked individually on initial tasks, students who had collaborated earlier did better on the mastery and transfer tasks. Previous studies have shown that students often do not support their claims with data (McNeill & Krajcik, 2007; Sandoval & Millwood, 2005). There are also reports that students often do not provide reasoning (warrants) for justifying their claims (Kuhn & Reiser, 2005; McNeill & Krajcik, 2007). However, in this PLGI context, students support most of their claims with data and warrants. This is a promising finding because it shows that students do provide evidence for their claims and can explain how that evidence connects to those claims.

Another significant finding here is that students infrequently provided backings as part of their arguments. This finding is consistent with Bell's (2000) work, in which middle school students rarely offered backings. Even though students in our study generally provided warrants, they did not often go beyond that step to validate the argument further. The ChemActivity curricular material used by students in our study contained scaffolds that promoted explanations. This structure is consistent with previous research showing that instructional frameworks with scaffolding promote student argumentation (Berland & Reiser, 2009). However, argumentation may be further enhanced by providing students with prompts based on the Toulmin's scheme itself (Chin & Osborne, 2010a, 2010b; Kaya, 2013; Weinberger, Stegmann, & Fischer, 2010) for eliciting data, warrants, and backings. This higher level of scaffolding would be likely to help students construct better and stronger arguments more frequently. How teachers implement argumentation in the classroom is an area that needs much further investigation (McNeill & Krajcik, 2008), especially since incorporation of argumentation into instructional practice has been found to be a slow and challenging process for teachers (Osborne, Simon, Christodoulou, Howell-Richardson, & Richardson, 2013). Our future studies entail examining peer leader intervention with student groups, including the role of different peer leader verbal behaviors (Gillies, 2004, 2006) on eliciting student arguments.

Finally, this study demonstrates that students were able to resolve their incorrect claims through argumentation without the peer leader providing them with the correct answer, or even intervening in any way. This is an important finding because many chemistry instructors may feel that they must provide students with correct answers in order for them to understand concepts and solve chemistry problems. This lends further credence to the effectiveness of cooperative learning environments where students work on their own without much direct assistance from the instructors (Slavin, 1990). In this PLGI setting, students were able to resolve almost all of the inaccurate claims by engaging in collective argumentation without peer leader intervention. Previous work has shown that engaging in argumentation could improve understanding of science concepts (Jiménez-Aleixandre et al., 2000; Mason, 1996; Zohar & Nemet, 2002) and improve reasoning skills (Mercer et al., 2004; Simon & Maloney, 2007). This study demonstrates that students were able to resolve inaccurate claims via argumentation, an indication of independent learninglearning.

We would like to thank Nicole Becker for her contribution to this work. This material is based on work supported by the National Science Foundation under Grant No. DUE-0618758. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Abi-El-Mona, I., & Abd-El-Khalick, F. (2011). Perceptions of the nature and 'goodness' of argument among college students, science teachers, and scientists. International Journal of Science Education, 33(4), 573–605.

Amigues, R. (1988). Peer interaction in solving physics problems: Sociocognitive confrontation and metacognitive aspects. Journal of Experimental Child Psychology, 45(1), 141–158.

Aydeniz, M., Pabuccu, A., Cetin, P. S., & Kaya, E. (2012). Argumentation and students' conceptual understanding of properties and behaviors of gases. International Journal of Science and Mathematics Education, 10(6), 1303–1324.

Barron, B. (2003). When smart groups fail. Journal of the Learning Sciences, 12(3), 307–359.

Bell, P. (2000). Scientific arguments as learning artifacts: Designing for learning from the web with KIE. International Journal of Science Education, 22(8), 797–817.

Becker, N., Rasmussen, C., Sweeney, G., Wawro, M., Towns, M., & Cole, R. (2013). Reasoning using particulate nature of matter: An example of a sociochemical norm in a university-level physical chemistry class. Chemistry Education Research and Practice, 14(1), 81–94.

Berland, L. K., & Hammer, D. (2012). Framing for scientific argumentation. Journal of Research in Science Teaching, 49(1), 68–94.

Berland, L. K., & Lee, V. R. (2012). In pursuit of consensus: Disagreement and legitimization during small-group argumentation. International Journal of Science Education, 34(12), 1857–1882.

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

Bianchini, J. A. (1997). Where knowledge construction, equity, and context intersect: Student learning of science in small groups. Journal of Research in Science Teaching, 34(10), 1039–1065.

Bianchini, J. A. (1999). From here to equity: The influence of status on student access to and understanding of science. Science Education, 83(5), 577–601.

Chang, S.-N., & Chiu, M.-H. (2008). Lakatos' scientific research programmes as a framework for analysing informal argumentation about socio-scientific issues. International Journal of Science Education, 30(13), 1753–1773.

Chen, J.-J., Lin, H.-S., Hsu, Y.-S., & Lee, H. (2011). Data and claim: The refinement of science fair work through argumentation. International Journal of Science Education, Part B, 1(2), 147–164.

Chin, C., & Osborne, J. (2010a). Students' questions and discursive interaction: Their impact on argumentation during collaborative group discussions in science. Journal of Research in Science Teaching, 47(7), 883–908.

Chin, C., & Osborne, J. (2010b). Supporting argumentation through students' questions: Case studies in science classrooms. Journal of the Learning Sciences, 19(2), 230–284.

Cohen, E. G. (1984). Talking and working together: Status, interaction and learning. In Peterson P., Wilkinson L. C., & Hallinan M. (Eds.), Instructional groups in the classroom: Organization and processes (pp. 171–187). Orlando, FL: Academic Press.

Cole, R., Becker, N., Towns, M., Sweeney, G., Wawro, M., & Rasmussen, C. (2012). Adapting a methodology from mathematics education research to chemistry education research: Documenting collective activity. International Journal of Science and Mathematics Education, 10(1), 193–211.

Daubenmire, P. L., & Bunce, D. M. (2008). What do students experience during POGIL instruction?. In Moog R. S. & Spencer J. N. (Eds.), Process-oriented guided inquiry learning. Washington, DC: Oxford University Press.

de Lima Tavares, M., Jiménez-Aleixandre, M. P., & Mortimer, E. F. (2010). Articulation of conceptual knowledge and argumentation practices by high school students in evolution problems. Science & Education, 19(6–8), 573–598.

Drane, D., Smith, H., Light, G., Pinto, L., & Swarat, S. (2005). The gateway science workshop program: Enhancing student performance and retention in the sciences through peer-facilitated discussion. Journal of Science Education and Technology, 14(3), 337–352.

Duschl, R. (2008). Quality argumentation and epistemic criteria. In Erduran S. & Jiménez-Aleixandre M. P. (Eds.), Argumentation in science education: Perspectives from classroom-based research (pp. 159–175). Dordrecht, Netherlands: Springer.

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

Enderle, P., Walker, J. P., Dorgan, C., Sampson, V. (2010, March 22). Assessment of scientific argumentation in the classroom: An observation protocol. Paper presented at the Annual International Conference for the National Association for Research in Science Teaching, Philadelphia, PA.

Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. Science Education, 88(6), 915–933.

Esmonde, I. (2009). Ideas and identities: Supporting equity in cooperative mathematics learning. Review of Educational Research, 79(2), 1008–1043.

Evagorou, M., & Osborne, J. (2013). Exploring young students' collaborative argumentation within a socioscientific issue. Journal of Research in Science Teaching, 50(2), 209–237.

Evagorou, M., Sadler, T. D., & Tal, T. (2011). Metalogue: Assessment, Audience, and Authenticity for Teaching SSI and Argumentation. In T. D. Sadler (Ed.), Socio-scientific issues in the classroom. Netherlands: Springer Science+Business Media.

Evagorou, M., Papanastasiou, E., Osborne J. (2011, September 7). The case of designing and validating a tool to assess 11–14 year old students written argumentation. Paper Presented at the European Science Education Research Association (ESERA), Lyon, France.

Evans, D. (1991). The realities of un-tracking a high school. Educational Leadership, 48(8), 16–17.

Foong, C.-C., & Daniel, E. G. S. (2011). Assessing students' arguments made in socio-scientific contexts: The considerations of structural complexity and the depth of content knowledge. Procedia—Social and Behavioral Sciences, 9, 1120–1127.

Furberg, A., & Arnseth, H. (2009). The importance of socio-cultural context for understanding students' meaning making in the study of genetics. Cultural Studies of Science Education, 4(1), 211–219.

Gillies, R. M. (2004). The effects of communication training on teachers' and students' verbal behaviours during cooperative learning. International Journal of Educational Research, 41(3), 257–279.

Gillies, R. M. (2006). Teachers' and students' verbal behaviours during cooperative and small-group learning. British Journal of Educational Psychology, 76, 271–287.

Gosser, D. K., Kampmeier, J. A., & Varma-Nelson, P. (2010). Peer-led team learning: 2008 James Flack Norris Award address. Journal of Chemical Education, 87(4), 374–380.

Gotwals, A. W., & Songer, N. B. (2010). Reasoning up and down a food chain: Using an assessment framework to investigate students' middle knowledge. Science Education, 94(2), 259–281.

Jiménez-Aleixandre, M. P., Rodrìguez, A. B., & Duschl, R. A. (2000). "Doing the lesson" or "doing science": Argument in high school genetics. Science Education, 84(6), 757–792.

Jiménez-Aleixandre, M. P., & Erduran, S. (2008). Argumentation in science education: An overview. In Jiménez-Aleixandre M. P. & Erduran S. (Eds.), Argumentation in science education: Perspectives from classroom-based research (pp. 3–28). Dordrecht, the Netherlands: Springer.

Johnson, D. W., Johnson, R. T., & Smith, K. A. (1998). Cooperative learning returns to college: What evidence is there that it works? Change, 30(4), 26–35.

Kaya, E. (2013). Argumentation practices in classroom: Pre-service teachers' conceptual understanding of chemical equilibrium. International Journal of Science Education, 35(7), 1139–1158.

Keith, W. M., & Beard, D. E. (2008). Toulmin's rhetorical logic: What's the warrant for warrants? Philosophy and Rhetoric, 41(1), 22–50.

Kelly, G. J., Druker, S., & Chen, C. (1998). Students' reasoning about electricity: Combining performance assessments with argumentation analysis. International Journal of Science Education, 20(7), 849–871.

King, A. (1992). Facilitating elaborative learning through guided student-generated questioning. Educational Psychologist, 27(1), 111–126.

King, A. (1998). Transactive peer tutoring: Distributing cognition and metacognition. Educational Psychology Review, 10(1), 57–74.

Kuhn, L., & Reiser, B. (2005). Students constructing and defending evidence-based scientific explanations. Paper presented at the Annual International Conference of the National Association for Research in Science Teaching, Dallas, TX.

Kulatunga, U., & Lewis, J. E. (2013). Exploration of peer leader verbal behaviors as they intervene with small groups in college general chemistry. Chemistry Education Research and Practice. Advance online publication. DOI: 10.1039/C3RP00081H

Lewis, S. E., & Lewis, J. E. (2005). Departing from Lectures: An evaluation of a peer-led guided inquiry alternative. Journal of Chemical Education, 82(1), 135–139.

Lewis, S. E., & Lewis, J. E. (2008). Seeking effectiveness and equity in a large college chemistry course: An HLM investigation of peer-led guided inquiry. Journal of Research in Science Teaching, 45(7), 794–811.

Lubben, F., Sadeck, M., Scholtz, Z., & Braund, M. (2009). Gauging students' untutored ability in argumentation about experimental data: A South African case study. International Journal of Science Education, 32(16), 2143–2166.

Mason, L. (1996). An analysis of children's construction of new knowledge through their use of reasoning and arguing in classroom discussions. Qualitative Studies in Education, 9, 411–433.

McNeill, K. L. (2011). Elementary students' views of explanation, argumentation, and evidence, and their abilities to construct arguments over the school year. Journal of Research in Science Teaching, 48(7), 793–823.

McNeill, K. L., & Krajcik, J., (2007) Middle school students' use of appropriate and inappropriate evidence in writing scientific explanations. Paper presented at the Carnegie Symposium on Cognition, Mahwah, NJ.

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

Mercer, N., Dawes, L., Wegerif, R., & Sams, C. (2004). Reasoning as a scientist: Ways of helping children to use language to learn science. British Education Research Journal, 30, 359–377.

Micari, M., Streitwieser, B., & Light, G. (2006). Undergraduates leading undergraduates: Peer facilitation in a science workshop program. Innovative Higher Education, 30(4), 269–286.

Mitchell, S. (1996). Improving the quality of argument in higher education: Interim report. London: School of Education Middlesex University.

Moog, R. S., & Farrell, J. J. (2008). Chemistry: A guided inquiry (2nd ed.). Hoboken, NJ: Wiley & Sons, Inc.

Naylor, S., Keogh, B., & Downing, B. (2007). Argumentation and primary science. Research in Science Education, 37(1), 17–39.

Nussbaum, E. M. (2008). Collaborative discourse, argumentation, and learning: Preface and literature review. Contemporary Educational Psychology, 33(1), 345–359.

Nussbaum, E. M. (2011). Argumentation, dialogue theory, and probability modeling: Alternative frameworks for argumentation research in education. Educational Psychologist, 46(2), 84–106.

Nussbaum, E. M., Sinatra, G. M., & Poliquin, A. (2008). Role of epistemic beliefs and scientific argumentation in science learning. International Journal of Science Education, 30(15), 1977–1999.

Osborne, J. (2010). Arguing to learn in science: The role of collaborative, critical discourse. Science, 328(5977), 463–466.

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. Journal of Research in Science Teaching, 41(10), 994–1020.

Osborne, J., Simon, S., Christodoulou, A., Howell-Richardson, C., & Richardson, K. (2013). Learning to argue: A study of four schools and their attempt to develop the use of argumentation as a common instructional practice and its impact on students. Journal of Research in Science Teaching, 50(3), 315–347.

Patton, M. Q. (2002). Qualitative research and evaluation methods. Thousand Oaks, CA: Sage Publishing.

Plantin, C. (2005). L'argumentation. Histoire, théories et perspectives. Paris: Presses Universitaires de France.

Quitadamo, I. J., Brahler, C. J., & Crouch, G. J. (2009). Peer-led team learning: A prospective method for increasing critical thinking in undergraduate science courses. Science Educator, 18(1), 29–39.

Rasmussen, C., & Stephan, M. (2008). A methodology for documenting collective activity. In Kelly A. E., Lesh R. A., & Baek J. Y. (Eds.), Handbook of innovative design research in science, technology, engineering, mathematics (STEM) education (pp. 195–215). New York: Taylor and Francis.

Reusser, K. (2001). Co-constructivism in educational theory and practice. In Smelser N. J., Baltes P. B., & Weiner F. E. (Eds.), International encyclopedia of the social and behavioral sciences. Oxford, UK: Pergamon/Elsevier Science.

Richmond, G., & Striley, J. (1996). Making meaning in classrooms: Social processes in small-group discourse and scientific knowledge building. Journal of Research in Science Teaching, 33(8), 839–858.

Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In O'Malley C. (Ed.), Computer-supported collaborative learning (pp. 69–197). Berlin: Springer.

Rozenszayn, R., & Ben-Zvi Assaraf, O. (2011). When collaborative learning meets nature: Collaborative learning as a meaningful learning tool in the ecology inquiry based project. Research in Science Education, 41(1), 123–146.

Sampson, V., & Clark, D. (2008). Assessment of the ways students generate arguments in science education: Current perspectives and recommendations for future directions. Science Education, 92(3), 447–472.

Sampson, V., & Clark, D. (2009). The impact of collaboration on the outcomes of scientific argumentation. Science Education, 93(3), 448–484.

Sampson, V., & Clark, D. (2011). A comparison of the collaborative scientific argumentation practices of two high and two low performing groups. Research in Science Education, 41(1), 63–97.

Sandoval, W. A., & Millwood, K. (2005). The quality of students' use of evidence inwritten scientific explanations. Cognition and Instruction, 23(1), 23–55.

Simon, S., & Maloney, J. (2007). Activities for promoting small-group discussion and argumentation. School Science Review, 88(324), 49–57.

Slavin, R. E. (1977). Classroom reward structure: An analytical and practical review. Review of Educational Research, 47(4), 633–650.

Slavin, R. E. (1990). Achievement effects of ability grouping in secondary schools: A best-evidence synthesis. Review of Educational Research, 60(3), 471–499.

Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. Contemporary Educational Psychology, 21, 46–69.

Sperry, R. A., & Tedford, P. (2008). Implementing Peer-Led Team Learning in introductory computer science courses. Journal of Computing in Small Colleges, 23(6), 30–35.

Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. Review of Educational Research, 69(1), 21–51.

Stephan, M., & Rasmussen, C. (2002). Classroom mathematical practices in differential equations. Journal of Mathematical Behavior, 21(4), 459–490.

Tien, L. T., Roth, V., & Kampmeier, J. A. (2002). Implementation of a Peer-Led Team Learning instructional approach in an undergraduate organic chemistry course. Journal of Research in Science Teaching, 39(7), 606–632.

Toulmin, S. (1958). The uses of argument. Cambridge: Cambridge University Press.

Veenman, S., Denessen, E., van den Akker, A., & van der Rijt, J. (2005). Effects of a cooperative learning program on the elaborations of students during help seeking and help giving. American Educational Research Journal, 42, 115–151.

Venville, G. J., & Dawson, V. M. (2010). The impact of a classroom intervention on grade 10 students' argumentation skills, informal reasoning, and conceptual understanding of science. Journal of Research in Science Teaching, 47(8), 952–977.

Webb, N. M. (1980). Group process and learning in an interacting group. The Quarterly Newsletter of the Laboratory of Comparative Human Cognition, 2, 10–15.

Webb, N. M. (1984). Sex differences in interaction and achievement in cooperative small groups. Journal of Educational Psychology, 76(1), 33–44.

Weber, K., Maher, C., Powell, A., & Lee, H. S. (2008). Learning opportunities from group discussions: Warrants become the objects of debate. Educational Studies in Mathematics, 68(3), 241–267.

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.

Wittrock, M. C. (1974). Learning as a generative process. Educational Psychologist, 11(2), 87–95.

Yu, S.-M., & Yore, L. (2012). Quality, evolution, and positional change of university students' argumentation patterns about organic agriculture during an argument-critique-argument experience. International Journal of Science and Mathematics Education, 1–22.

Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. Journal of Research in Science Teaching, 39(1), 35–62.