International Journal of Science Education
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tsed20

Personally-Seeded Discussions to Scaffold Online Argumentation

Douglas B. Clark [a] & Victor D. Sampson [a]
[a] Arizona State University, USA

To cite this article: Douglas B. Clark & Victor D. Sampson (2007): Personally-Seeded Discussions to Scaffold Online Argumentation, International Journal of Science Education, 29:3, 253-277

To link to this article: http://dx.doi.org/10.1080/09500690600560944

International Journal of Science Education
Vol. 29, No. 3, 26 February 2007, pp. 253–277

RESEARCH REPORT

Personally-Seeded Discussions to Scaffold Online Argumentation

Douglas B. Clark* and Victor D. Sampson
Arizona State University, USA

Research shows that scientific knowledge develops through a process of decision-making as well as discovery, and that argumentation is a genre of discourse crucial to the practice of science. Students should therefore be supported in understanding the scientific practices of dialectical and rhetorical argumentation as part of learning about scientific inquiry. This study focuses on supporting scientific argumentation in the classroom through a customized online discourse system. “Personally-seeded discussions” support learning and collaboration through an activity structure that elicits, shares, and contrasts students’ own ideas to engage them in the discourse of science argumentation and inquiry. Students use an online interface to build principles to describe data they have collected. These principles become the seed comments for the online discussions. The software sorts students into discussion groups with students who have built different principles so that each discussion group can consider and critique multiple perspectives. This study explores the efficacy of this personally-seeded approach based on a coding scheme developed by Erduran, Osborne, and Simon that analyzes argument structure from a Toulmin perspective. As part of this exploration, the study outlines a method for parsing personally-seeded discussions into oppositional episodes for analysis, and discusses future directions for supporting argumentation in asynchronous online discussions.

Keywords: *Dialectical Argumentation; Rhetorical Argumentation; Scientific Inquiry*

Introduction

Scientific knowledge develops through a process of a decision-making as well as discovery (Kuhn, 1970; Kuhn, Shaw, & Felton, 1997; Latour & Woolgar, 1986). Dialectical argumentation and rhetorical argumentation are genres of discourse critical to this process (Driver, Newton, & Osborne, 2000; Kuhn, 1993; Latour & Woolgar, 1986; Lemke, 1990; Longino, 1994; Siegel, 1995; Toulmin, 1958). In order to promote student understanding of the nature and process of science, students must participate in discourse that parallels the practices of scientific

---

*Corresponding author. Arizona State University, College of Education, Division of Curriculum and Instruction, Payne Hall, PO Box 870911, Tempe, AZ 85287-0911, USA. Email: dbc@asu.edu

ISSN 0950-0693 (print)/ISSN 1464-5289 (online)/07/030253–25

DOI: 10.1080/09500690600560944

communities (Lemke, 1990; Rosebery, Warren, & Conant, 1992; Schauble, Glaser, Duschl, Schulze, & John, 1995). Taken together, these findings demonstrate the importance of developing students' epistemic understanding of the social and cultural scientific practices of dialectical and rhetorical argumentation through classroom-based inquiry activities.

This study analyzes a customized online discourse system designed to integrate and support scientific argumentation within the classroom. The context is an online thermal equilibrium inquiry laboratory designed for eighth-grade students. The students use a special interface to build principles to describe the data they collect in the laboratory portion of the project. These principles become the seed comments for the online discussion. The software sorts the students into discussion groups with students who have built different principles so that each discussion group represents multiple perspectives. Students then follow a set of guidelines to critique one another's principles. By having students explain and defend their own principles, students take interest in responding to and critiquing the other ideas in the discussion. This structure allows students to engage in the discourse of scientific argumentation and inquiry in a way that parallels scientific communities.

We explore the efficacy of this personally-seeded discussion approach using a coding scheme developed by Erduran, Osborne, and Simon (2004; Simon, Osborne, & Erduran, 2003) that analyzes argument structure from a Toulmin perspective. In the process, we present a method for parsing this type of asynchronous online discussion into "oppositional episodes" for analysis of argument structure. In conclusion, we discuss future directions to support argumentation in asynchronous online discussions.

The Role of Argumentation in Science and Science Education

There has been a great deal of interest and effort in recent years in reforming science education to reflect more accurately the practice of the scientific community by engaging students in authentic scientific inquiry. Textbooks often define scientific inquiry as a process of asking questions, developing a means to collect data in order to answer those questions, interpreting the data, and drawing conclusions that can be used to contribute to our understanding of the world. Yet this view of the process of inquiry neglects the socially situated nature of science (Kuhn, 1970; Latour & Woolgar, 1986; Suppe, 1998). From a social perspective, the acceptance of any claim by the scientific community does not necessarily rely on the inherent truth-value of the claim; rather, it relies on the degree to which others can be persuaded to accept the claim. This persuasion factor is due to the fact that the audience, and not the individual, is responsible for judging the validity and quality of any theoretical claim. Scientific reasoning should therefore be understood not only as a process of inference, but also as a process of persuasion.

Understanding the persuasion factor in the scientific process has important implications for science education. Argumentation is central to analyzing data and information, writing persuasive explanations, and engaging in direct dialog. Argumentation is therefore a social and collaborative process that is necessary to solve problems and advance knowledge (Duschl & Osborne, 2002) rather than a competition between individuals (Solomon, 1998). Argumentation includes any dialog that addresses "the coordination of evidence and theory to support or refute an explanatory conclusion, model, or prediction" (Osborne, Erduran, & Simon, 2004, p. 995). When defined in this manner, scientific argumentation is part of a broader social practice that is used to persuade other people at the heart of scientific inquiry (Duschl & Osborne, 2002).

Supporting Argumentation in Science Classrooms

The integration of argumentation into science curricula is currently upheld as a core requirement for a successful science program (Kuhn, 1993; Driver et al., 2000; Duschl & Osborne, 2002). In order to participate in inquiry and argumentation, students need to learn how to make sense of arguments and develop an understanding of the social and cultural scientific practices that produce them. Driver et al. argue that "to know science is to know not only what a phenomenon is but also how it relates to other events, why it is important, and how this particular view of the world came to be" (2000, p. 297). From this perspective, teaching science as a process of discovery is not enough. Students need to understand the discursive modes of argumentation that are valued by the discipline and the practices that are used by the discipline in order to construct knowledge (Driver et al., 2000; Newton, Driver, & Osborne, 1999). To develop this type of understanding, students must participate in educational environments that parallel the social contexts of scientific communities (Herrenkohl, Palincsar, DeWater, & Kawasaki, 1999; Kuhn, 1993; Lemke, 1990; Rosebery et al., 1992). In other words, teaching science as a process of inquiry without giving students the opportunity to engage in argumentation in a social context fails to represent a core component of the nature of science or to establish a medium for the development of students' conceptual understanding of science (Duschl & Osborne, 2002).

Successful integration of argumentation into classroom activities depends upon the member of the classroom engaging in methods of discussion that parallel those of scientific communities and arriving at a common perspective on the phenomena discussed (Forman, Larreamendy-Joerns, Stein, & Brown, 1998). Unfortunately, a major barrier to the development of students' scientific argumentation skills is the lack of opportunity to participate in these types of activities within current pedagogical practices.

Research investigating whole-class discourse in science classrooms shows that the discussions taking place in science classrooms are dominated by a teacher-led structure. They often focus on factual recall and follow a pattern that has been described as "teacher initiation, student response, and teacher evaluation" (Goldman, Duschl, Ellenbogen, Williams, & Tzou, 2002, p. 259). Teachers use this type of interaction in the classroom to "develop and rehearse points which the teacher deems important," and to "determine whether or not the pupils can reproduce the answers they [the teachers] have in mind" (Newton et al., 1999, p. 563). This type of structure contributes to students learning facts, but it does not function well when the goal of instruction is to promote reasoning skills or to learn about the process of science (Driver et al., 2000). In these contexts, students are commonly given less than a second to formulate their answers (Rowe, 1974), which as a result typically consist of a single word or a short phrase rather than a reasoned argument containing an extended student contribution to the construction of a dialogic argument. In addition, teachers often sustain inequitable class participation, calling on boys more often than girls in science classes (Hsi & Hoadley, 1997). Fundamentally, according to Duschl & Osborne (2002), "such dialogue has limited educational value and maintains the power relations that support and structure classroom life." (p. 42). Thus the teacher, who seeks to establish the consensually-agreed scientific world view with the student, minimizes the opportunities for dialogic discourse (Duschl & Osborne, 2002) and ensures that the student neither controls nor even understands the focus of the discourse (Driver et al., 2000).

Small group discussions can also constrain student participation. When students work in groups, these opportunities are rarely organized in a way that encourages substantive discussion of the science involved. Instead, students tend to focus on the procedural aspects of the work (Alexopoulou & Driver, 1996; Driver et al., 2000) or they attempt to complete the task individually rather than collaboratively (Cohen, 1994). In addition, studies suggest that social interactions among peers who are engaged in small group discussions are strongly influenced by existing group dynamics. Students can actually regress in their scientific reasoning because of their perception of their abilities in relation to those of other group members (Alexopoulou & Driver, 1996). Furthermore, students who are influenced by social pressure may actually take on the views of their peers without understanding the other viewpoint or revising their personal view (Hsi & Hoadley, 1997).

In contrast to these ineffective attempts at classroom participation, research suggests that argumentation can be nurtured if both the epistemological and social discourse structures in the classroom are taken into account when designing inquiry activities. Studies suggest that the discursive modes by which scientific information and knowledge are communicated and represented are critical in fostering argumentation in the classroom. For example, Duschl and Osborne (2002) suggest that the minimum requirements for an argumentation-supportive context include the consideration of plural accounts of phenomena and a classroom community that sees all members as equal. Newton et al. (1999) argue that students need to participate actively through talk and writing in making sense of the scientific events, experiments, and explanations to which they are being introduced. Finally, and perhaps most importantly, the epistemic goals for classroom learning must emphasize "how we know what we know, and why we believe the beliefs of science to be superior or more fruitful than competing viewpoints" (Osborne, 2001, p. 43). This research suggests that argumentation can be embedded into inquiry-based activities if specific structures are in place within the classroom.

In summary, researchers have identified several requirements for supporting argumentation in science classrooms: (a) students must engage with plural accounts of phenomena and evidence to support multiple points of view (Bloom, 2001; Duschl & Osborne, 2002), (b) the learning environment must provide a context that fosters dialogic discourse, (c) tasks and activities given to groups must require collaboration in order to promote discourse between students (Cohen, 1994; Eichinger, Anderson, Palincsar, & David, 1991), (d) students must have enough time to understand the central concepts and underlying principles (Duschl & Osborne, 2002; Rowe, 1974), and (e) the teacher or learning environment must facilitate student-to-student talk without the limitations and rigidities characteristic of most teacher–student interactions (Bloom, 2001). These characteristics should be taken into account when creating learning environments to support scientific argumentation.

Computer-based Supports for Argumentation: Personally-seeded discussions

Structured environments have been built to support scientific argumentation, discourse, and knowledge refinement. Some of these environments such as Collaboratory Notebook (Edelson, Pea, & Gomez, 1996), CaMILE (Guzdial, Turns, Rappin, & Carlson, 1995), and Knowledge Forum/CSILE (Scardamalia, Bereiter, & Lamon, 1994) are learning environments unto themselves that focus heavily on knowledge collection and building. Others, such as SpeakEasy (Hoadley, Hsi, & Berman, 1995), Sensemaker (Bell, 1997), and the BGUILE data reporting section (Tabak, Smith, Sandoval, & Reiser, 1996), are part of larger inquiry environments. SpeakEasy focuses on sharing and discussing ideas. Sensemaker supports students sorting and classifying evidence for the purposes of argument creation for a larger debate with other students in a class. BGUILE aids students in connecting data and warrants to arguments for presentation to a class. In addition to these specialized environments, basic online threaded asynchronous forums in which discussions are held have also been shown to be effective in supporting classroom-based discourse (Collison, Elbaum, Haavind, & Tinker, 2000; Salmon, 2000). In these forums, students are able to use the technology to share their ideas with one another and to receive feedback by working as a group to build their ideas.

Whereas the online environments detailed focus either on sharing information or on preparing arguments for presentation, personally-seeded discussions focus specifically on engineering and supporting scientific argumentation within classroom discourse. The environment supports students in both the social discourse patterns as well as the epistemological goals of argumentation. Personally-seeded discussions (a) help students synthesize a principle to describe data that they have collected or found in light of other evidence from their classroom and homes, (b) create groups of students who have created different principles to describe the data, (c) facilitate online discourse among the students where they critique each other's principles in light of the evidence and work toward consensus through scientific argumentation based on the evidence, and (d) provide students with models of productive scientific argumentation.

The personally-seeded discussion system analyzed in this study was piloted in an online inquiry project. The project uses Web-based Integrated Science Environment

Internet software with custom simulation modeling, electronic peer critique, and laboratory components integrated to support students as they investigate thermal equilibrium (Figure 1). Students begin the inquiry project by making predictions about the temperature of everyday objects around them in the classroom. Students then use thermal probes to investigate the temperature of these objects and construct principles to describe the patterns they encounter. This first segment of the project attempts to cue students' conflicting ideas, including students' sense that objects are different temperatures because "they feel that way" and students' sense that objects in the room should be the temperature of the room because "what would make them be a different temperature?" In the second segment of the project, the software places students in electronic discussion groups with students who have constructed different principles explaining the data. Finally, in the third segment, the students experiment with a visualization designed to help students revise their experiential ideas related to their understanding of thermal equilibrium (Clark & Jorde, 2004).

The structural quality of the students' argumentation within the second segment of the project is the focus of the current study. After collecting their data, students create principles to describe patterns in the data. Research on students' initial conceptions about heat and temperature (Clark, 2000, 2001; Lewis, 1996; Linn & Hsi, 2000) identifies principles students typically use to describe heat flow and thermal equilibrium. This conceptual change research and earlier thermodynamics curriculum development (Lewis, Stern, & Linn, 1993) form the foundation of a new Web-based principle-builder interface that allows students to construct scientific principles from a set of predefined phrases and elements (Figure 2a). After students create their principles, the project software places the students in electronic discussion groups with students who have constructed different explanatory principles. A

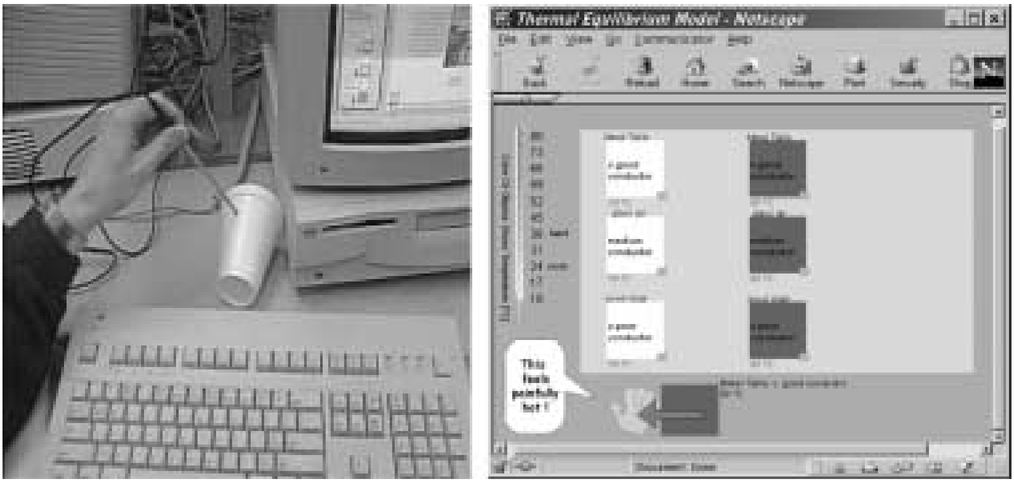
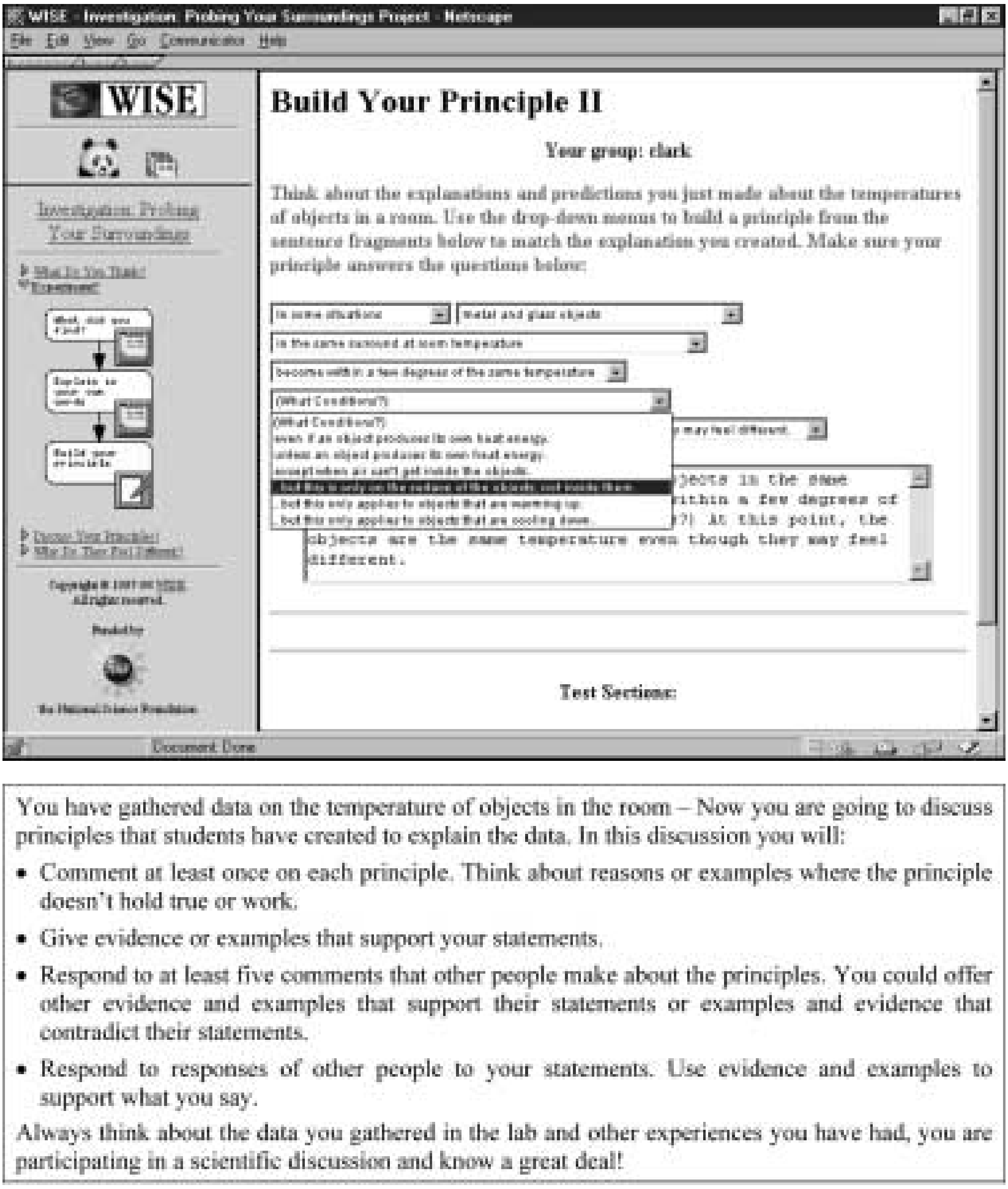


Figure 1. Students gather data, build principles, share ideas in a threaded online forum, and work with simulations

screenshot of a portion of an asynchronous discussion within the thermodynamics project from this study is included in Figure 3.

The student-constructed principles appear as the seed comments in the discussions. The discussions develop around the different perspectives represented in the



Build Your Principle II

Your group: clark

Think about the explanations and predictions you just made about the temperatures of objects in a room. Use the drop-down menus to build a principle from the sentence fragments below to match the explanation you created. Make sure your principle answers the questions below:

In some situations | metal and glass objects
in the same surround at room temperature
becomes within a few degrees of the same temperature
(What Conditions?)

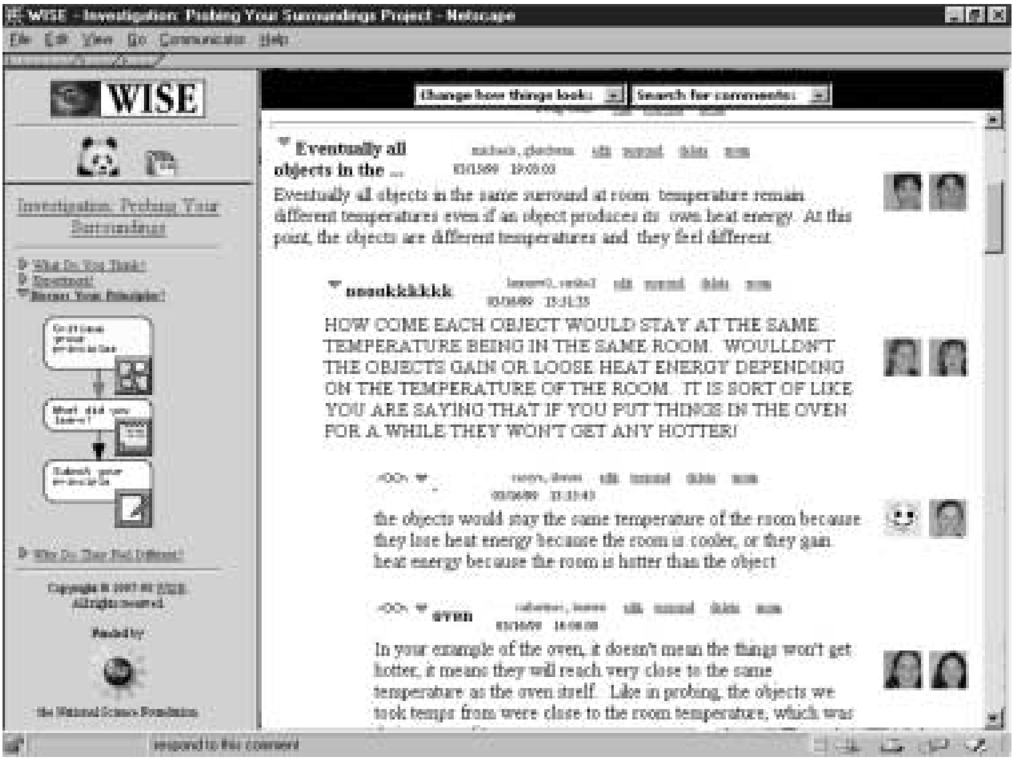
...jects in the same ...ithin a few degrees of ...?) At this point, the objects are the same temperature even though they may feel different.

You have gathered data on the temperature of objects in the room – Now you are going to discuss principles that students have created to explain the data. In this discussion you will:

- Comment at least once on each principle. Think about reasons or examples where the principle doesn't hold true or work.
- Give evidence or examples that support your statements.
- Respond to at least five comments that other people make about the principles. You could offer other evidence and examples that support their statements or examples and evidence that contradict their statements.
- Respond to responses of other people to your statements. Use evidence and examples to support what you say.

Always think about the data you gathered in the lab and other experiences you have had, you are participating in a scientific discussion and know a great deal!

Figure 2. (a) Students use the principle-builder interface to construct scientific principles that become initial discussion comments. (b) Instructions are included at the top of each online discussion



Eventually all objects in the ...

Eventually all objects in the same surround at room temperature remain different temperatures even if an object produces its own heat energy. At this point, the objects are different temperatures and they feel different.

ooookkkkkk

HOW COME EACH OBJECT WOULD STAY AT THE SAME TEMPERATURE BEING IN THE SAME ROOM. WOULLDN'T THE OBJECTS GAIN OR LOOSE HEAT ENERGY DEPENDING ON THE TEMPERATURE OF THE ROOM. IT IS SORT OF LIKE YOU ARE SAYING THAT IF YOU PUT THINGS IN THE OVEN FOR A WHILE THEY WON'T GET ANY HOTTER!

the objects would stay the same temperature of the room because they lose heat energy because the room is cooler, or they gain heat energy because the room is hotter than the object

green

In your example of the oven, it doesn't mean the things won't get hotter, it means they will reach very close to the same temperature as the oven itself. Like in probing, the objects we took temps from were close to the room temperature, which was

Figure 3. Typical discussion generated from one initial seed claim

seed comments, ideally through a process of comparison, clarification, and justification. As part of this process, the students are required to support their assertions and claims with evidence from their laboratories and other experiences (see student instructions in Figure 2b). This process attempts to elicit self-explanation by helping students focus other students' attention on possible inconsistencies in their explanations and on reasoning, plausibility, completeness, and other attributes of "good explanations." In these discussions, all students and their ideas become critical resources with the common goal of refining individual student ideas.

Analysis of Student Argumentation

To make judgments about argumentation quality, researchers over the past decade have developed methods to identify the essential features of an argument. These methods have been used to examine the structure of student arguments in conversation (Forman et al., 1998; Kelly, Druker, & Chen, 1998; Resnick, Salmon, Zeitz, Wathen, & Holowchak, 1993) and in writing (Bell & Linn, 2000; Kelly & Takao, 2002). To date, most of these investigations of student discourse rely on Toulmin's (1958) model for argument structure in one way or another. In these studies, emphasis is placed on the identification of the structural features of arguments (e.g., claims, data, warrants, backings, and qualifiers) and the process of argumentation, especially in terms of how students provide warrants for claims. Such approaches seek to identify the absence or presence of the components of argument and use this information to assess argumentation quality. Structural analyses of student arguments contribute to our understanding of how students assimilate the desired practices of argumentation (Driver et al., 2000) and provide a great deal of information about the form and type of reasoning that students use when they construct arguments based on their everyday experiences (Simon et al., 2003).

For example, the analysis by Jimenez-Aleixandre, Rodrigues, and Duschl (2000) of high school students' discussions about genetics found that arguments constructed and co-constructed by the students tend to rely heavily on claims that lack backings and warrants. Also, in these group discussions, students tend to use a large number of opposing claims rather than rebuttals (Jimenez-Aleixandre et al., 2000). Kelly et al. applied Toulmin's framework to spoken arguments produced by pairs of students as they participated in a problem-solving activity about electrical circuits. Similarly, they found that students do not usually provide warrants for their claims unless they are challenged (Kelly et al., 1998). In addition, Zeidler (1997) found that students are likely to affirm a claim if they believe a premise is true rather than false despite warrants contrary to their beliefs, and are often unsure of what constitutes convincing evidence. Taken together, the conclusions from these and other studies (Coleman, 1998; Driver et al., 2000; Ohlsson, 1992; Voss & Means, 1991) suggest that students' everyday discourse strategies are often inappropriate when students attempt (a) to make inferences or draw conclusions based on evidence, (b) to construct and articulate coherent explanations, or (c) to respond to and evaluate the explanations made by others.

Although we have learned a great deal about the nature of the arguments that students construct, assessing the overall quality of an extended written argument or dialogic argument has been problematic. Several researchers (Driver et al., 2000; Erduran et al., 2004; Kelly et al., 1998) report difficulty making objective distinctions between the various Toulmin components when analyzing dialectic argumentation. For example, identifying a student's statement within a dialogic argument as a claim, warrant, or backing can be difficult since the identification of these components often requires context analysis. In addition, because students often share in the construction of an argument rather than defending one point of view over another, some researchers have found that many of the statements made by students cannot be classified using Toulmin's framework. Jimenez-Aleixandre et al. (2000) suggest that a structural analysis of dialogic argumentation needs to include additional argumentative operations (such as clarification, query, or support for another's claim).

Erduran, Osborne, and Simon's Coding of Structural Argumentation Quality

The coding scheme of Erduran et al. (2004; Simon et al., 2003) focuses on the quality of dialogic argumentation within whole-class discussions. In order to assess

Table 1. Argumentation hierarchy from Erduran et al. (2004) and Simon et al. (2003)

Level 5 Argumentation displays an extended argument with more than one rebuttal with grounds
Level 4 Argumentation shows arguments with a claim with a clearly identifiable rebuttal with grounds
Level 3 Argumentation has arguments with a series of claims or counter claims with grounds and the occasional weak rebuttal
Level 2 argumentation has arguments consisting of claims with grounds but No rebuttals
Level 1 Argumentation consists of arguments that are a simple claim versus a counter claim or a claim versus a claim

the quality of this type of argumentation, Erduran et al. focus solely on argument structure and the identification of the structural components used by students as they co-construct a dialogic argument. To analyze an argument using Erduran et al.'s hierarchy, the elements of an argument must first be identified. These first elements include the extent to which students make use of claims, opposing claims, rebuttals, and grounds (based on Toulmin's model). Erduran et al. (2004; Simon et al., 2003) collapse Toulmin's data, warrants, and backings into a single "grounds" code due to the practical difficulties of reliably differentiating among these argumentation components. Once these elements are identified, the overall quality of the argument can be ranked according to its level of structural sophistication using a five-level hierarchy that is outlined in Table 1.

The ranking of argument quality is based on how often the students incorporate the elements (i.e., claims, grounds, and rebuttals) of the coding scheme into their argument. Their coding scheme focuses specifically on the incorporation of rebuttals supported by grounds within an argument. We choose Erduran et al.'s coding scheme for the current study (a) because it is compatible with the discourse format within the personally seeded discussion system, and (b) because using the coding scheme of another research group facilitates benchmarking the efficacy of the personally-seeded discussion approach.

Research Questions

Early pilot studies suggest that students contribute more comments in personally-seeded discussions than in standard online discussions with generic principles as seeds (Clark, 2004; Cuthbert, Clark, & Linn, 2002). By having students explain and defend their own principles, students may not only take an interest in their own ideas, but also take interest in responding to and critiquing the other ideas in the discussion due to an increase in social relevance (Hoadley, 1999; Hoadley & Linn, 2000). The current study asks two questions about the nature and quality of the argumentation supported:

1. What is the structural nature of the argumentation within personally-seeded discussions?

2. Can personally-seeded discussions scaffold high levels of scientific argumentation as defined by Erduran, Osborne, and Simon?

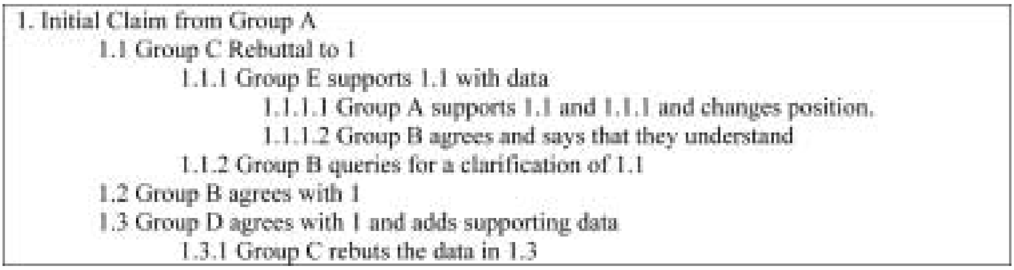
Methods

Participants

Eight online discussions involving a total of 84 students have been randomly chosen from four classes of eighth-grade students who completed the project during one semester under the supervision of an experienced teacher who has worked extensively with the researchers. The public school is located in a diverse city and has an even distribution of boys and girls. The classes are typical eighth-grade physical science classes, labeled neither "honors" nor "remedial." Each online discussion involves approximately five pairs of students. Students work on the project in pairs over the course of six class periods (5 hr in total). The discussions begin at the start of the fourth class period and extend through the end of the fifth class period. To represent multiple perspectives, the software assigns student pairs to discussions with students who have created different principles, as discussed above.

Asynchronous Online Discussions and Units of Analysis: Oppositional episodes

The discussions are threaded and asynchronous. That means that the students may respond to any contribution in the discussion at any time. As is typical in asynchronous threaded forums, response comments are placed by the software underneath the comment to which they are replying (the parent comment) and indented. A structural outline of a hypothetical fragment from a typical discussion generated from one initial seed claim is outlined in Figure 4 to help explain the coding of an "episode" as well as to explain the structure of asynchronous online discussions for readers not familiar with these forums. Note that because the discussion in Figure 4 is asynchronous and threaded, the children comments to parent Comment 1.1 (i.e., Comments 1.1.1, 1.1.1.1, 1.1.1.2, and 1.1.2) may have been contributed before or after Comment 1.2 was contributed. Structurally, we know only that later comments



1. Initial Claim from Group A
    - 1.1 Group C Rebuttal to 1
        - 1.1.1 Group E supports 1.1 with data
            - 1.1.1.1 Group A supports 1.1 and 1.1.1 and changes position.
            - 1.1.1.2 Group B agrees and says that they understand
        - 1.1.2 Group B queries for a clarification of 1.1
    - 1.2 Group B agrees with 1
    - 1.3 Group D agrees with 1 and adds supporting data
        - 1.3.1 Group C rebuts the data in 1.3

Figure 4. Structural outline of a hypothetical fragment from a typical discussion generated from one initial seed claim to help explain the coding of an "episode" as well as to explain the structure of asynchronous online discussions for readers not familiar with these forums

at a given depth are sequential (e.g., Comment 1.1 preceded 1.2, which preceded 1.3). Within the actual discussions, a time stamp accompanies each comment to establish the precise time of contribution.

Erduran et al. (2004; Simon et al., 2003), search through the transcripts looking for "oppositional episodes" between students to code for argumentation quality. The current study defines a discourse episode to include an initial claim, one of that initial claim's second-tier comments, and all the children of that specific second-tier comment. In Figure 4, there are therefore three episodes defined by Comments 1.1, 1.2, and 1.3. The 1.1 episode includes Comments 1, 1.1, 1.1.1, 1.1.1.1, 1.1.1.2, and 1.1.2. The 1.2 episode includes Comments 1 and 1.2 only. The 1.3 episode includes Comments 1, 1.3, and 1.3.1. Each of these three episodes is analyzed as a potential oppositional episode. In the Figure 4 example, the episodes defined by Comments 1.1 and 1.3 contain opposition and are therefore included in the analysis using Erduran et al.'s hierarchy. The 1.2 episode does not include opposition and therefore is not ranked under Erduran et al's hierarchical scheme but is coded for inclusion in other analyses.

Base Coding Scheme and Hierarchy for Rating Argumentation Structure

This study uses the three base codes for students' comments from Erduran et al.'s scheme: (1) claim, (2) grounds, and (3) rebuttal. We also add four additional base codes to further characterize the epistemic operations of student communications and interactions in the online environment: (4) support, (5) query, (6) emotive appeal, and (7) off-task comments. Our coding scheme assigns these base codes in the following manner.

1. **Claim.** The initial principle that each student group creates through the principle maker interface is placed in the discussion as that group's initial seed comment. This initial principle for each student pair is considered the initial claim for arguments that develop based on that claim. Student pairs may also introduce opposing claims within an episode in the online discussions (e.g., "All the objects will remain different temperatures in the same surroundings").
2. **Grounds.** As discussed earlier, grounds include data, warrants, and backings (e.g., "The metal chair felt different but it was room temperature in our experiment"). Erduran et al. (2004; Simon et al., 2003) collapsed this category because of pragmatic challenges in reliably differentiating data, warrants, and backings in student transcripts. We did not classify a comment as including grounds if it simply restates data, warrants, or backings used in the claim to which it is replying or referring.
3. **Rebuttal.** Rebuttals include attacks on the grounds of a claim or attacks directly on a claim (e.g., highlighting a portion of a claim that is invalid). We diverge from Erduran et al. on this point of the scheme. Erduran et al. focus specifically and exclusively on attacks on the grounds of a claim, where we also include attacks directly on a claim. For example, students can provide evidence from a

laboratory activity that precludes acceptance of a portion of a claim ("You wrote that the temp would be the same even if the objects produced heat. But they won't be the same look at your graph for the computer and the table!"). For students in these discussions, we feel providing evidence that directly contradicts a portion of a claim is an important component of rebutting a claim. Toulmin (1958) defines a rebuttal as both the "circumstances in which the general authority of the warrant would have to be set aside" and the "exceptional conditions which might be capable of defeating or rebutting the warranted conclusion." We therefore categorize statements as a rebuttal if any portion of that statement directly attacks the grounds of a claim or the claim itself.

4. **Support.** Support comprises comments that support earlier claims or rebuttals. These may be useful additions including grounds for the support (e.g., "I agree with Joseph because when we did our experiment with the probes the screen and the chair leg did not change 8 degrees which Nancy said") or may simply involve a comment announcing agreement with a claim or rebuttal (e.g., "We agree with you because our principle is the same as yours"). Sometimes, the support comment involves a clarification of wording or intent for an earlier comment from the group (e.g., "I was concentrating on metal because it was the best conductor, and the rest were good insulators"). Occasionally, a support comment announces agreement with a rebuttal or claim and signifies a change of position for the contributors from an earlier stated claim or rebuttal.
5. **Query.** Queries comprise comments that ask for clarification of earlier comments (e.g., "What do you mean when you say …?") or inquire where other groups stand on an issue (e.g., "Do we all agree …?"). Queries do not include comments that are used to question the validity of another group's claim or grounds, which are classified as rebuttals (e.g., "How do you know that those two objects are really different temperatures?").
6. **Emotive Appeal.** Occasionally comments are not part of rational argumentation and are primarily emotional in content but are focused on veracity and authority (e.g., "Don't agree with him! He's wrong!").
7. **Off Task.** Comments also occasionally veer off task (e.g., "Nice haircut, John!").

All comments are assigned a code in light of the comment to which they reply, which means that the comments are coded in context rather than as individual statements. Inter-rater reliability using this coding system is high and is discussed in detail in the Results and Discussion section. Using these base codes, student arguments are then rated using Erduran et al.'s hierarchy of argumentation structure (see Table 1).

Figure 5 provides an example of the coding process. In this example, there are two potential episodes defined by the two Level 2 comments (i.e., the first episode includes Comments 1 and 1.1, and the second episode includes Comments 1, 1.2, and 1.2.1). The first episode in Figure 5 (involving Comments 1 and 1.1) does not involve opposition and is therefore not considered an oppositional episode for ranking in Erduran et al.'s argumentation hierarchy. The individual comments for this

1 Claim	***Group 1*** *(This initial statement by student pair 1 is the principle they created to describe their lab data. These principles became the seed comments for the online threaded discussion.)* Immediately all objects in the same surround at room temperature become within a few degrees of the same temperature unless an object produces its own heat energy. At this point the objects are within a few degrees even though they may feel different.
1.1 Support Grounds	***Group 2*** *(response to Group 1)* We agree with this because in the lab we observed that almost all objects were around the same temperature.
1.2 Rebuttal Grounds	***Group 3*** *(response to Group 1)* How do you know that the temperature changes immediately? Wouldn't it change at different rates depending on how good a conductor the object is? Couldn't it reach the temperature at a slower or faster rate, although it will eventually reach the same or close to the same temperature of the room?
1.2.1 Support	***Group 4*** *(response to Group 3)* You're right!!! The materials' ability to conduct heat will determine how fast it will heat up.

Figure 5. Coding for the comments associated with one initial seed claim

non-oppositional episode are still coded with the base codes, however, for other analyses in this study. The second episode (involving Comments 1, 1.2, and 1.2.1) does include opposition and is therefore considered an oppositional episode. Differentiating between rebuttals and opposing claims is the challenging part of the coding scheme. The coding scheme considers a rebuttal to be a comment that directly and explicitly contradicts part of a comment. In Figure 5, Group 3's comment is coded as a rebuttal to the "Immediately" part of Group 1's initial claim. Group 3 supports its rebuttal with data about conductivity affecting the rate of temperature change. This example qualifies as a Level 4 argument in the Erduran et al. hierarchy outlined in Table 1.

Results and Discussion

The results from the eight discussions are organized by argumentation score in Table 2. The eight discussions include 122 total episodes comprising 416 student comments. Of these episodes, 63 qualify as oppositional episodes and 59 do not.

Non-oppositional Episodes

Most non-oppositional episodes tend to be very short. The mean number of comments within a non-oppositional episode is 2.37 and the mean depth is 2.25 comments. (Depth is the longest chain of comments within an episode.) In most non-oppositional episodes, students do not include grounds for their support statements. For example, in the two non-oppositional episodes defined by Comments 2.2

Table 2. Comment code results by episode category

				Opposing claims		Rebuttals		Support						
Category of contribution	Number of comments	Mean episode depth	Initial claims	Without grounds	With grounds	Without grounds	With grounds	Without grounds	With grounds	Change position	Clarification	Queries	Emotional appeals	Off-task elements
All episodes														
Total	416		122	19	2	24	80	53	25	8	14	30	25	36
Mean	3.41	2.73		0.16	0.02	0.2	0.66	0.43	0.2	0.07	0.11	0.25	0.2	0.3
Non-oppositional episodes														
Total	140		59	0	0	0	0	37	16	2	1	7	7	17
Mean	2.37	2.25		0	0	0	0	0.63	0.27	0.03	0.02	0.12	0.12	0.29
Oppositional episodes														
Total	276		63	19	2	24	80	16	9	6	13	23	18	19
Mean	4.38	3.17		0.3	0.03	0.38	1.27	0.25	0.14	0.1	0.21	0.37	0.29	0.3
Oppositional episodes Rated 1														
Total	23		8	10	0	0	0	0	0	0	1	4	2	1
Mean	2.88	2.63		1.25	0	0	0	0	0	0	0.13	0.5	0.25	0.13
Oppositional episodes Rated 2														
Total	0		0	0	0	0	0	0	0	0	0	0	0	0
Mean	0	0		0	0	0	0	0	0	0	0	0	0	0
Oppositional episodes Rated 3														
Total	30		12	0	0	13	0	1	1	0	2	3	0	0
Mean	2.5	2.42		0	0	1.08	0	0.08	0.08	0	0.17	0.25	0	0
Oppositional episodes Rated 4														
Total	77		25	3	0	0	25	4	5	4	5	8	4	0
Mean	3.08	2.88		0.12	0	0	1	0.16	0.2	0.16	0.2	0.32	0.16	0
Oppositional episodes Rated 5														
Total	146		18	6	2	11	55	11	3	2	5	8	12	18
Mean	8.11	4.33		0.33	0.11	0.61	3.06	0.61	0.17	0.11	0.28	0.44	0.67	1

and 2.3 in Figure 6, the students simply state agreement and repeat the portion of the comment with which they agree. In the 16 support statements that do include grounds, the students typically explain why particular aspects of the initial claim would be true (see Table 2). An example of this is illustrated by the following Group A2 comment:

> **Group A1** [This initial statement is the seed claim by Group A1]
> In some situations some objects in the same surround at room temperature remain different temperatures even if an object produces its own heat energy, At this point, the objects are different temperatures even though they may feel the same. **(Claim)**
>
> **Group A2** [Response to Group A1]
> It's true because some objects have more heat energy than others and are different temperatures. It is the material the objects are made out of that determines the temperatures. **(Support with Grounds)**
>
> **Group A3** [Response to Group A2]
> We kind of agree with what you are saying. Like [Tina] said the wording is confusing, but you are right with what you are saying with the temperatures of the items. **(Support without Grounds)**

Support statements without grounds comprise 37 of the 55 supporting comments in the non-oppositional episodes. The remaining two support instances involve students acknowledge changing their position in line with the initial claims. An example of this support with change of position is now shown:

> **Group B1** [This initial statement is the seed claim by Group B1]
> Eventually all objects in the same surround at room temperature become within a few degrees of the same temperature unless an object produces its own heat energy. At this point, the objects are within a few degrees even though they may feel different. **(Claim)**
>
> **Group B2** [Response to Group B1]
> We agree with you more than we agree with what we put because now when we look back at what we put. Eventually it does get within a few degrees of each other. **(Support – Change of Claim)**

In summary, the non-oppositional episodes tend to be relatively unsophisticated in terms of scientific discourse structures. Students tend to accept what is written in the claim and move onward.

Oppositional Episodes

Oppositional episodes are the focus of Erduran et al.'s analysis. We therefore present oppositional episodes by argumentation score as well as in aggregate. The curricular structure of the students' project focuses on critiquing the principles of the other students. As an apparent result of this structure, we see very few simple "claim versus opposing claim" episodes (argumentation Levels 1 or 2 in Table 1). The few episodes (8 of the 63 oppositional episodes) that do meet these criteria are brief (mean comments = 2.88) and involve no data, warrants, or backings (i.e., grounds). An example of a Level 1 argument is as follows:

> **Group C1** [This initial statement is the seed claim by Group C1]
> Eventually all objects in the same surround at room temperature become the same temperature even if an object produces its own heat energy. At this point the objects are within a few degrees even though they may feel the same. **(Claim)**
>
> **Group C2** [Response to Group C1]
> I don't agree with you because they will never all reach the same temperature. **(Counter-Claim)**
>
> **Group C3** [Response to Group C2]
> They all won't reach the same temp but they will be sorts close. **(Support without Grounds)**

Level 3 arguments, which include weak rebuttals, comprise 12 of the 63 oppositional episodes. These weak episodes are also brief (mean comments = 2.5). An example of a Level 3 argument is now shown:

> **Group D1** [This initial statement is the seed claim by Group D1]
> Eventually all objects in the same surround at room temperature become within a few degrees of the same temperature except when air can't get inside the objects. At this point, the objects are different temperatures even though they may feel the same. **(Claim)**
>
> **Group D2** [Response to Group D1]
> I disagree about the air going in the objects part because I don't think it matters. **(Rebuttal without Grounds)**

Episodes involving rebuttals with grounds (Level 4) comprise 25 of the 63 oppositional episodes and are longer (mean comments = 3.08). The 18 episodes involving multiple sequential rebuttals (Level 5) have a mean number of comments of 8.11. This much longer average is heavily weighted by the single longest episode, which spanned 26 comments. Overall, the oppositional episodes include many more instances of clarification (13 versus 1) and queries (23 versus 7) than non-oppositional queries, even as a percentage of total comments. An example of a Level 5 argument can be seen in the following:

> **Group E1** [This initial statement is the seed claim by Group E1]
> Eventually all objects in the same surround at all temperatures become within a few degrees of the same temperature, but this is only on the surface of the objects, not inside them. At this point, the objects are within a few degrees even though they may feel different. **(Claim)**
>
> **Group E2** [Response to Group E1]
> Actually, in our lab, when we compared temperatures of the metal chair and a computer screen, the temperature had a relatively drastic difference. (8 degrees…ooh! big difference!) According to your principle, the computer screen and the metal chair should eventually become within just a few degrees difference of temperature. However, that cannot be right. It just does not seem possible, does it? **(Rebuttal with Grounds)**
>
> **Group E3** [Response to Group E2]
> When we did our experiment the only difference with the metal chair leg is that it changed by only about 1 degree. **(Rebuttal with Grounds)**
>
> **Group E1** [Response to Group E3]

> I agree with Joseph because when we did our experiment with the probes the screen and the chair leg did not change 8 degrees which Nancy said. I think she should check her solution again and not say I am wrong. **(Support with Grounds)**
>
> **Group E4** [Response to Group E2]
> I think also if you have two objects and you leave them in a room I think they will get to room temperature. The only thing I am saying is that the two objects will probably not reach the same temperature at the same time. This might happen if you have a good conductor that will heat up very fast because heat energy can pass through it easily. If the second object was an insulator like wood it would not let heat pass through it very easily so it would take longer for it to heat up. It also matters what was the starting temperature of the objects when they were put in the room. This would change if the object would come to room temperature in a day or in months. **(Rebuttal with Grounds)**
>
> **Group E1** [Response to Group E2]
> You could be right about drastic changes in the temperature of the computer screen and the metal chair leg, but in our observations of the metal leg of the chair and the computer screen did not have an 8 degree difference in temperature. I think you had your computer on to long and it heated it up. So we would have different solutions. **(Rebuttal with Grounds)**

In their research, Simon et al. (2003) found that, during face-to-face student discussions, 32% of the oppositional episodes include clearly identifiable rebuttals (Level 4 or 5) while the majority of the oppositional episodes involve arguments that consist of claims with grounds without rebuttals. In the personally-seeded discussions, 68% of the oppositional episodes in this study classify as Level 4 or Level 5 arguments. As discussed in the following Caveats in Comparing Results section, these numbers are not directly comparable because of a variation in the coding schemes, but they do suggest that personally-seeded discussions scaffold high structural levels of scientific argumentation. In addition to argumentation quality, personally-seeded discussions also seem worthwhile in terms of the 2% baseline quantity of typical classroom discourse that involves argumentation (Simon et al., 2003). In comparison with this baseline for quantity of argumentation in classrooms, personally-seeded discussions scaffold significant levels of argumentation with minimal teacher training (i.e., 276 of 416 comments occur in oppositional episodes and 80 of 416 comments involve rebuttals with grounds).

An interesting dichotomy, however, involves the issue of scientific content accuracy. Many of the rebuttals with grounds used by the students involved real-world examples or data from the project or prior laboratories. For example:

> Why do you disagree? Remember when we did the potato lab? Even the things that were well insulated would eventually reach room temperature. Also remember how circuits use copper wire because we know that copper is a good conductor! Obviously the good conductor would reach room temperature first as opposed to bad conductors.

From the perspective of argumentation, rebuttals like these are excellent in their use of data and warrants. The students in the previous example, however, raise important data from their prior laboratory but confuse electrical and thermal conductivity in their real-world example. Normative conceptual content unfortunately does not always accompany desirable argumentation structure. One significant challenge for future development of personally-seeded discussions is to create supports that prevent persuasive students from leading other students astray with solid argumentation backed by non-normative grounds.

Variation among Discussions

A pedagogical challenge in designing personally-seeded discussions involves the variation in character and tenor of discussions. Each of the discussions evidences a different character depending on the tone set by the participants. Statistically, analysis of variance shows no significant difference among discussions or among class periods in terms of argumentation scores or number of comments. Several other factors, however, such as acceptable behavior, formats, tactics, and rhetoric, appear to be informally negotiated among the members of the discussion. Off-task behavior and emotional appeals tend to occur in clusters, for instance. For example, only 4 of 36 off-task comments occur in isolation within an episode. Similarly, use of evidence seems also to come in clusters. For example, only 25 of 80 instances where rebuttals are backed with grounds occur in isolation within an episode. In the personally-seeded discussions of this study, students are insulated from the other discussions groups. Students are therefore exposed only to the small community in which they are participating. While personally-seeded discussions orchestrate the social setting effectively in terms of argumentation structure overall, further work needs to be done to fine tune the social interactions to optimize the discourse. Future work will need to focus on moving students away from negative social and epistemic practices while facilitating the dissemination of exemplary social and epistemic practices.

Caveat in Comparing Results with Erduran, Osborne, and Simon

The hierarchy of argumentation structure developed and used by Erduran et al. centers around the idea that only arguments that rebut the grounds of a person's argument can undermine the beliefs of that individual. In other words, oppositional episodes that do not rebut the grounds have no potential to change the thinking of the participants because the basis of each participant's beliefs rests on the grounds used as justification. This definition of a rebuttal seems appropriate for debates steeped in social values (e.g., the "socio-scientific" debates in Erduran et al.'s curriculum about whether zoos are good or bad). Erduran et al. describe socio-scientific debates based on Osborne's, Erduran's, & Simon's definition (2004) as debates "where social practices are constantly examined and reformed in light of scientific evidence" (p. 996). In socio-scientific debates, attacking a grounded claim (e.g., "zoos are good because people can see the animals and want to protect them") with a grounded reply (e.g., "zoos are bad because the animals are unhappy") is often a counter-claim rather than a rebuttal. The attack presents another perspective but does not disqualify the initial claim, and therefore fits with Erduran et al.'s coding definition that only comments that attack grounds can be coded as rebuttals.

Our study focuses on debates that Erduran et al. would term “scientific” that require empirical argumentation concerning the concept of thermal equilibrium. We define a rebuttal to include direct attacks on a portion of the original claim. This definition is appropriate in an empirical context because grounds can be provided to fully refute the original claim. For example, a claim that “objects stay different temperatures even if you leave them out on the table for a long time because I’ve felt them and they feel different” can be rebutted by saying that “the objects actually become the same temperature like when we did the lab and the temperatures of the wood table and the bottle of soda both became 23 degrees after a long time” or by saying that “the objects only feel different even though they are the same temperature because they have different thermal conductivities.” From our perspective, both the first reply attacking the claim and the second reply attacking the grounds constitute rebuttals of the initial claim that the “objects stay different temperatures.”

Because our definition of a rebuttal includes these attacks directly on the original claim in addition to attacks on the grounds supporting the original claim, however, our version of the coding scheme results in an elevation in the ranking of some of the episodes. We acknowledge Erduran et al.’s rationale for coding social debates but assert that our definition correctly values the epistemic value of attacking a portion of the claim directly in this type of debate, particularly when accompanied by appropriate grounds. We have discussed this issue with Jonathon Osborne (Osborne & Clark, personal communication) of Erduran et al. (2004) in person, but further work will be required to refine the value and quality codings and ratings for the valid epistemic moves that students make in argumentation. Regardless, because of this difference in coding definitions, we do not intend for the scores to be directly compared in terms of which curriculum resulted in “higher” or “lower” scores. Rather, the scores are compared qualitatively simply to suggest that the personally-seeded discussions result in successful levels of argumentation, particularly in light of the scientific context, which Simon et al. (2003) and Osborne et al. (2004) found to be more challenging for students than socio-scientific contexts.

Appropriateness of Parsing Method

From a methodological standpoint, the practice of defining episodes based on the second-level comments seems appropriate given the average number of comments and the average depth of comment chains within the episodes in this study. The difference between mean number and mean depth is about 0.1 for non-oppositional episodes and oppositional episodes of Level 3 or less. The difference for Level 4 arguments is 0.2. These small differences between depth (length of longest chain in episode) and number of comments (total comments in episode) suggest that episodes tend to be linear rather than branched. This linear quality suggests that the current study’s parsing scheme seldom combines significantly branched discussion as a single episode. Only Level 5 the episodes approach a grey area, where the mean number of comments is 8.11 and the mean depth is 4.33 for the 18 episodes. One of these episodes is unusually large in comparison with the others, containing 26 comments on its own and a depth of 7 with multiple branches. It might be more appropriate to subdivide episodes of this size into multiple episodes representing each of the substantial branches within the large episode. Further analysis of extended episodes will be required to resolve this particular issue. Overall, however, this issue applied to only one episode out of 122 and the parsing method for defining discourse episodes proved appropriate for the vast majority of the episodes.

Inter-rater Reliability of the Coding Scheme

The coding scheme proposed in this paper seems fairly robust. The eight discussions were initially coded by the lead author. The second author was later trained on the coding and parsing scheme and coded the eight discussions again. This resulted in an inter-rater reliability of 91%. The largest category of differences between the two coders involved distinguishing off-task comments versus emotive appeals. By refining that definition in the coding scheme, inter-rater reliability climbed to 94%. The next two largest categories of disagreement involved nine instances where one author assigned the “support” code where the other author assigned the “rebuttal” code, and nine instances where one author assigned the “rebuttal with grounds” code where the other author assigned the “rebuttal without grounds” code.

These instances were resolved through discussion, but whereas the distinction between off-task and emotive appeal became immediately apparent, we believe that disagreements about the presence or lack of grounds will continue to involve a small percentage of gray areas in coding. The disagreements between support and rebuttal codes involved differences in perspectives about which comment a reply primarily addressed. If a comment follows a claim and a rebuttal and supports the rebuttal, it may also contain elements that directly rebut the initial claim. Our revised definitions clarify that if a comment a comment should be coded primarily with respect to its direct parent comment. To address the “grounds” versus “no grounds” issue, we revised our definitions to clarify that restating grounds previously included in the episode does not qualify as adding grounds. Upon resolving these two coding issues, our inter-rater agreement increased to 98%. The remaining 2% of differences between the two coders involved situations where one coder assigned a code and the other coder did not think that the situation was clear enough to warrant a code being assigned. These differences were resolved through discussion for this study, but in terms of the coding scheme itself this last 2% of variance between coders seems relatively inevitable but acceptable.

Conclusions

Simon et al. (2003) found that during face-to-face student discussions 32% of the oppositional episodes include clearly identifiable rebuttals (Level 4 or Level 5), while the majority of the oppositional episodes involve arguments that consist of claims with grounds without rebuttals. In the personally-seeded discussions, 68% of the oppositional episodes in this study classify as Level 4 or Level 5 arguments. As discussed in the earlier Caveats in Comparing Results section, these numbers are not directly comparable because of a variation in the coding schemes, but they do suggest that personally-seeded discussions scaffold high structural levels of scientific argumentation. In addition to argumentation quality, personally-seeded discussions also seem worthwhile in terms of the 2% baseline quantity of typical classroom discourse that involves argumentation (Simon et al., 2003). In comparison with this baseline for quantity of argumentation in classrooms, personally-seeded discussions scaffold significant levels of argumentation with minimal teacher training (i.e., 276 of 416 comments occur in oppositional episodes and 80 of 416 comments involve rebuttals with grounds).

Simon et al.'s findings suggest that altering teacher practice is more critical than student training in terms of changing the quality of argumentation in the classroom. In their study, the important variable is the teacher's experience with the interventions. Importing curricular change into the classroom might be more easily accomplished with technology. Future inquiry might focus on the degree to which extended interaction with personally-seeded discussions changes teacher or student argumentation behavior in the classroom when outside of the online scaffolds. What would be the impact over time of participating in these online discussions in terms of changing student practice? Will students internalize any of these improved argumentation strategies and use them unprompted in classroom settings?

One critical issue under investigation involves content. This study focuses on the ability of personally-seeded discussions to scaffold argumentation structure as measured by Erduran et al.'s hierarchical ranking scheme. The results of the study suggest that personally-seeded discussions are successful by this measure. The Results and Discussion section, however, also provides evidence demonstrating that desirable argumentation structure is not always accompanied by normative conceptual science content. The current phase of our research is investigating the interaction of science content and argumentation structure in personally-seeded discussions.

This paper continues the discussion about creating effective environments to support science inquiry and argumentation. While in-class inquiry discourse typically involves only a small percentage of the students and marginalizes many of the other class members, text-based environments offer the possibility of supporting a much broader range of students (Hsi & Hoadley, 1997). Text-based collaborative environments offer a natural choice because they allow students to participate directly in the linguistic medium of scientific discourse while engaging in inquiry and argumentation. If discourse is important to science, then the opportunity to interact with the actual medium and process of scientific discourse is exceptionally valuable. The results of this study suggest that carefully structured online environments can effectively scaffold student participation in this scientific discourse.

Acknowledgements

This research was supported by the National Science Foundation through grant number 0334199 (Technology Enhanced Learning in Science). The authors would like to acknowledge and thank Doris Jorde and the Computer as Learning Partner

Group for their collaboration and help with the data collection, curriculum development, and research out of which this paper has grown.

References

Alexopoulou, E., & Driver, R. (1996). Small-group discussion in physics: Peer interaction modes in pairs and fours. *Journal of Research in Science Teaching, 33*(10), 1099–1114.

Bell, P. (1997). *Using argument representations to make thinking visible for individuals in groups.* Paper presented at the Computer Supported Clever to Learning Conference '97, University of Toronto, Canada.

Bell, P., & Linn, M. C. (2000). Scientific arguments as learning artifacts: Designing for learning from the web with KIE. *International Journal of Science Education, 22*(8), 797–818.

Bloom, J. W. (2001). Discourse, cognition, and chaotic systems: An examination of students' argument about density. *The Journal of the Learning Sciences, 10*(4), 447–492.

Clark, D. B. (2000). *Scaffolding knowledge integration through curricular depth.* Unpublished doctoral dissertation, University of California, Berkeley, CA.

Clark, D. B. (2001). *New representations of student knowledge integration in CLP: Theories or repertoires of ideas?* Paper presented at the AERA, Seattle, WA.

Clark, D. B. (2004). Hands-on labs in internet environments: Teaching thermal equilibrium. In M. C. Linn, E. A. Davis, & P. Bell (Eds.), *Internet environments for science education.* Mahwah, NJ: Lawrence Erlbaum Associates.

Clark, D. B., & Jorde, D. (2004). Helping students revise disruptive experientially-supported ideas about thermodynamics: Computer visualizations and tactile models. *Journal of Research in Science Teaching, 41*(1), 1–23.

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research, 64,* 1–35.

Coleman, E., B. (1998). Using explanatory knowledge during collaborative problem solving in science. *The Journal of the Learning Sciences, 7*(3&4), 387–427.

Collison, G., Elbaum, B., Haavind, S., & Tinker, R. (2000). *Facilitating online learning: Effective strategies for moderators.* Madison, WI: Atwood Publishing.

Cuthbert, A., Clark, D. B., & Linn, M. C. (2002). Wise learning communities: Design considerations. In K. A. Renninger & W. Shumar (Eds.), *Building virtual communities: Learning and change in cyberspace* (pp. 215–246). Cambridge, UK: Cambridge University Press.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education, 84*(3), 287–213.

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

Edelson, D. C., Pea, R. D., & Gomez, L. (1996). The collaboratory notebook: Support for collaborative inquiry. *Communications of the ACM, 39,* 32–33.

Eichinger, D., Anderson, C. W., Palincsar, A. S., & David, Y. M. (1991, April). *An illustration of the roles of content knowledge, scientific argument, and social norm in collaborative problem solving.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Erduran, S., Osborne, J., & Simon, S. (2004). The role of argument in developing scientific literacy. In K. Boersma, O. deJong, H. Eijkelhof, & M. Goedhart (Eds.), *Research and the quality of science education.* Dordrecht, The Netherlands: Kluwer Academic Publishers.

Forman, E., A., Larreamendy-Joerns, J., Stein, M., K., & Brown, C. A. (1998). "You're going to want to find out which and prove it": Collective argumentation in a mathematics classroom. *Learning and Instruction, 8*(6), 527–548.

Goldman, S. R., Duschl, R. A., Ellenbogen, K., Williams, S. M., & Tzou, C. (2002). Science inquiry in a digital age: Possibilities for making thinking visible. In V. Oostendorp (Ed.), *Cognition in a digital age* (pp. 253–281). Mahwah, NJ: Erlbaum.

Guzdial, M., Turns, J., Rappin, N., & Carlson, D. (1995). Collaborative support for learning in complex domains. In J. L. Schanse & E. L. Cunnius (Eds.), *Proceedings of CSCL '95: Computer supported collaborative learning* (pp. 157–160). Hillsdale, NJ: Lawrence Erlbaum Associates.

Herrenkohl, L., Palincsar, A., DeWater, L., & Kawasaki, K. (1999). Developing scientific communities in classrooms: A sociocognitive approach. *The Journal of Learning Sciences, 8*(3&4), 451–494.

Hoadley, C. (1999). *Scaffolding scientific discussion using socially relevant representations in networked multimedia.* Unpublished doctoral dissertation, University of California, Berkeley, CA.

Hoadley, C. M., Hsi, S., & Berman, B. P. (1995). The multimedia forum kiosk and speakeasy. In *Proceedings of ACM multimedia '95* (pp. 363–364). New York: ACM Press.

Hoadley, C., & Linn, M. C. (2000). Teaching science through on-line peer discussions: Speakeasy in the knowledge integration environment. *International Journal of Science Education, 22,* 839–857.

Hsi, S., & Hoadley, C. M. (1997). Productive discussion in science: Gender equity through electronic discourse. *Journal of Science Education and Technology, 6*(1), 23–36.

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

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.

Kelly, G. J., & Takao, A. (2002). Epistemic levels in argument: An analysis of university oceanography students' use of evidence in writing. *Science Education, 86,* 314–342.

Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. *Science Education, 77*(3), 319–337.

Kuhn, D., Shaw, V., & Felton, M. (1997). Effects of dyadic interaction on argumentative reasoning. *Cognition and Instruction, 15*(3), 287–315.

Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.

Latour, B., & Woolgar, S. (1986). *Laboratory life: The construction of scientific facts.* Princeton, NJ: Princeton University Press.

Lemke, J. (1990). *Talking science: Language, learning, and values.* Norwood, NJ: Ablex.

Lewis, E. L. (1996). Conceptual change among middle school students studying elementary thermodynamics. *Journal of Science Education and Technology, 5*(1), 3–31.

Lewis, E. L., Stern, J., & Linn, M. C. (1993). The effect of computer simulations on introductory thermodynamics understanding. *Educational Technology, 33*(1), 45–58.

Linn, M. C., & Hsi, S. (2000). *Computers, teachers, peers: Science learning partners.* Mahwah, NJ: Lawrence Erlbaum Associates.

Longino, H. (1994). The fate of knowledge in social theories of science. In F. F. Schmidt (Ed.), *Socializing epistemology: The social dimension of knowledge.* Lanham, MD: Rowman & Littlefield.

Monk, M., & Osborne, J. (1997). Placing the history and philosophy of science in the curriculum: A model for the development of pedagogy. *Science Education, 81*(4), 405–424.

Newton, P., Driver, R., & Osborne, J. (1999). The place of argumentation in the pedagogy of school science. *International Journal of Science Education, 21*(5), 553–576.

Ohlsson, S. (1992). The cognitive skill of theory articulation: A neglected aspect of science education? *Science & Education, 1,* 181–192.

Osborne, J. (2001). Promoting argument in the science classroom: A rhetorical perspective. *Canadian Journal of Science, Mathematics, and Technology Education, 1*(3), 271–290.

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

Resnick, L., B., Salmon, M., Zeitz, C., M., Wathen, S., H., & Holowchak, M. (1993). Reasoning in conversation. *Cognition and Instruction, 11*(3&4), 347–364.

Rosebery, A. S., Warren, B., & Conant, F. R. (1992). Appropriating scientific discourse: Findings from language minority classrooms. *The Journal of the Learning Sciences, 2*(1), 61–94.

Rowe, M. (1974). Relation of wait time and the rewards to the development of language, logic and fate control: Part ii—rewards. *Journal of Research in Science Teaching, 11*(4), 291–308.

Salmon, G. (2000). *E-moderating: The key to teaching and learning online.* London & Sterling, VA: Kogan Page.

Scardamalia, M., Bereiter, C., & Lamon, M. (1994). The CSILE project: Trying to bring the classroom into world 3. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 201–228). Cambridge, MA: MIT Press.

Schauble, L., Glaser, R., Duschl, R., Schulze, S., & John, J. (1995). Students' understanding of the objectives and procedures of experimentation in the science classroom. *Journal of the Learning Sciences, 4*(2), 131–166.

Siegel, H. (1995). Why should educators care about argumentation? *Informal Logic, 17*(2), 159–176.

Simon, S., Erduran, S., & Osborne, J. (2002, April). *Enhancing the quality of argumentation in school science.* Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, New Orleans, LA.

Simon, S., Osborne, J., & Erduran, S. (2003). Systemic teacher development to enhance the use of argumentation in school science activities. In J. Wallace & J. Loughran (Eds.), *Leadership and professional development in science education: New possibilities for enhancing teacher learning* (pp. 198–217). London & New York: Routledge Falmer.

Solomon, J. (1998). About argument and discussion. *School Science Review, 80*(291), 57–62.

Suppe, F. (1998). The structure of a scientific paper. *Philosophy of Science, 65*(3), 381–405.

Tabak, I., Smith, B. K., Sandoval, W. A., & Reiser, B. J. (1996). Combining general and domain-specific strategic support for biological inquiry. In C. Frasson, G. Gauthier, & A. Lesgold (Eds.), *Proceedings of the Third International Conference on Intelligent Tutoring Systems (its '96)*. Montreal, Canada & New York: Springer-Verlag.

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

Voss, J., F., & Means, M., L. (1991). Learning to reason via instruction in argumentation. *Learning and Instruction, 1,* 337–350.

Zeidler, D., L. (1997). The central role of fallacious thinking in science education. *Science Education, 81,* 483–496.