

# Astronomy Education Review

Volume 4, Mar 2005 - Oct 2005

Issue 1

## **Promoting Argumentative Discourse: A Design-Based Implementation and Refinement of an Astronomy Multimedia Curriculum, Assessment Model, and Learning Environment**

by **Gita Taasoobshirazi**

University of Georgia

**Daniel T. Hickey**

University of Georgia

Received: 06/06/05, Revised: 08/05/05, Posted: 09/13/05

The Astronomy Education Review, Issue 1, Volume 4:53-70, 2005

© 2005, Gita Taasoobshirazi. Copyright assigned to the Association of Universities for Research in Astronomy, Inc.

### **Abstract**

This study uses an existing astronomy software program to encourage inquiry-based learning and argumentation while also ensuring gains on conventional classroom assessments and high-stakes achievement tests. To supplement the software program, a curriculum and three levels of assessment—quizzes, exam, and test—were designed and administered to 11th- and 12th- grade students during a four-week implementation. Students worked in groups for the quizzes and exam to engage in argumentative discourse. Because of the difficulties in engaging students in quality argumentation, the curriculum, assessments, and overall learning environment were examined for ways to further support argumentative discourse.

---

*Astronomy Village: Investigating the Universe* is a software program developed by NASA's Classroom of the Future (COTF). It is designed to engage students in inquiry-based learning and scientific argumentation through authentic investigations and simulations. In this regard, the software is representative of quite a few multimedia science programs developed in the last decade (see Note 1). To supplement the already-existing software package, a curriculum and three levels of assessment—quizzes, an exam that focused on the curriculum, and a test that targeted the required science standards—were designed and administered to students during a 20-hour, four-week implementation. Students worked in groups during two levels of the assessments—quizzes and exam—to engage in argumentative discourse. Because of the difficulties in engaging students in quality argumentation, the implementation was

examined for ways to improve the curriculum, assessments, and learning environment to further support argumentative discourse. Three main questions guided this study: (1) What is argumentation? (2) Why is it important? (3) How can students' argumentation skills be developed? This article first considers the prior research on the value of scientific argumentation and the challenges involved in supporting it. Furthermore, descriptions of the design-based methods guiding this study, the *Astronomy Village* learning environment, and the curriculum and assessment materials are provided. Then, an analysis of the argumentation of two groups of students in the classroom and a description of the overall classroom learning gains on the exam and test are provided. Finally, the lessons learned indicate that support for further teacher facilitation, development of a more detailed description of steps to help support and structure student argumentation, development of video rubric examples, adjustment of the quizzes and exam to the students' academic level, formation of groups of optimal size, and a curriculum that can help students gain the needed background knowledge would help students improve their argumentative discourse.

---

## 1. INTRODUCTION

For over a decade now, researchers have recognized the importance of the role of conversation and discourse when engaging students in science (Lemke 1990; Sutton 1992). Furthermore, in past years, researchers have focused on argumentative discourse in science education (Driver et al. 1994; Newton, Driver, & Osborne 1999; Simon, Erduran, & Osborne 2002). Despite these efforts, authentic argumentative discourse is rarely enacted in science classrooms (Driver, Newton, & Osborne 2000). Instead, science is presented as a collection of facts that students read and memorize. In the best-case circumstances, students perform experiments in which they follow directions to confirm already-learned understandings (Rudolph & Stewart 1998). More specifically, in most classrooms, a positivist view of science is prominent, and science is viewed as a subject in which there are clear right answers that emerge from uncontroversial data (Driver et al. 2000). However, this method of learning science is outmoded because it is authoritarian and not authentic.

Contemporary theorists insist that knowledge develops through social negotiation and through the judgment of the practicality of the ideas of others (Savery & Duffy 1996; Zhu 1998). In science, our knowledge is based on current theories; facts are facts only because there is widespread interpretation and agreement on the ideas rather than the understanding of some ultimate truth. It was once thought that the Earth was flat and that the Sun revolved around the Earth. Over time, theories and ideas change, and it is through scientific argumentation that understandings are tested to find the most viable explanations (Driver et al. 2000). Therefore, science should not involve transmitting a set of known facts to students. Asking students just to memorize facts and concepts prevents them from experiencing authentic science and gives them a false impression of how science is actually practiced. In addition to being unrealistic, it is also an authoritarian way of teaching science; "to ask of other human beings that they just accept and memorize what the science teacher says, without any concern for the meaning and justification of what is said, is to treat those human beings with disrespect and is to show insufficient care for their welfare" (Norris 1997, 252).

Engaging students in argumentation is expected to allow them to see that science is an ongoing process in which ideas are determined, questioned, and often changed or revised (Diehl 2000). Scientific argumentation, which involves "proposing, supporting, criticizing, evaluating, and refining ideas, some of which may conflict or compete, about a scientific subject" (Shin & McGee 2003, 1), engages students in

using evidence and theory to support or refute ideas or claims (Simon et al. 2002). Practicing scientists work together to communicate and evaluate their ideas and findings. In doing so, they justify or abandon their theories and beliefs with supporting or contradictory information gained through argumentation (Kuhn 1993). It is through this practice that ideas are questioned and changed. In the same way, students must be able to do more than just make claims or state facts; they should be able to support their claims using evidence to justify their ideas as the most viable.

To help describe the constitutive elements of argumentation, Toulmin (1958) proposed a general model that is used by researchers and teachers to help describe, develop, and evaluate argumentation (Driver et al. 2000; Jimenez-Aleixandre, Bugallo-Rodriguez, & Duschl 1997). The major components in Toulmin's model include:

- Claim: a conclusion, hypothesis, or opinion
- Data: facts that support the claim
- Warrants: proposed reasons using rules, theories, principles, and so on, to describe how the data support the claim
- Backings: commonly agreed assumptions that help justify warrants
- Rebuttals: specifying cases in which the claim will not or cannot be true

As students make claims, use data to support their claims, warrant their claims using data, and further justify their claims using backings and rebuttals, they not only learn how to engage in the authentic practices science, but they also learn the science content (American Association for the Advancement of Science 1993). Through argumentation, students become producers, not just consumers, of scientific knowledge (Brown & Campione 1998). As students argue and hear diverse ideas, they are able to reflect on their own ideas and the ideas of others, which can help them tackle misconceptions and develop better understandings. In fact, in a science class, the teacher may find that if the students are not discussing the science concepts, they are not learning the material (Cobern 1993). Furthermore, in addition to learning the content, through this structured discourse, students are able to become familiar with the way that scientific language is used (Lemke 1990), allowing their discourse to become more like that of practicing scientists.

## **2. DIFFICULTIES IN ENGAGING STUDENTS IN ARGUMENTATION**

Despite the evidence suggesting the importance of engaging students in argumentation, it is very difficult to do so. One of the biggest deterrents in engaging students in argumentation is the lack of resources, skills, and preparation that teachers face when trying to support argumentation in their classrooms (Driver et al. 2000). This most likely contributes to the lack of opportunity that students have to engage in discussion, and as a result, to the difficulties that students face when trying to participate in quality argumentation. However, because argumentation does not come naturally to people (Kuhn 1991), students need to be immersed in contexts that allow them to practice their argumentative skills.

Any effort to support argumentation raises numerous questions for which researchers and educators need guidance. One example is the issue of group size (Alexopoulou & Driver 1996). Teachers are often unsure of the group size that will allow students to participate in optimal discussion. When engaging in discussion, groups must be large enough to expose students to a diversity of opinions, but small enough

for all members of the group to participate. No set number has been found to support the best discussions, but researchers have suggested group sizes from three to six students (McClelland 1983; Slavin 1995).

Another factor that needs to be considered when engaging students in argumentation is the students' prior knowledge. A lack of prior knowledge of a subject or theory constrains students' ability to reason and thus warrant their claims using data (Koslowski 1996). Therefore, students are likely to feel and to be more capable of arguing about a topic when they have a certain degree of knowledge about that subject. The problem lies in the fact that students need to know something about science in order to engage in argumentation, but they must engage in argumentation to learn the needed content. However, learning both argumentation skills and content knowledge at the same time may be overwhelming. The question remains as to how to approach this concern.

### **3. DESIGN-BASED RESEARCH**

Central to this study's efforts to advance scientific argumentation is the idea of design-based research. Design-based research involves forming claims about learning, taking one's ideas and materials into a real and active classroom, and observing how they are enacted in order to improve both the theory and the implementation. In trying to describe design-based research more clearly, it is useful to point out some of the differences between traditional psychological methods and those of design-based research. Unlike traditional psychological experimentation, design-based research involves working exclusively in naturalistic settings rather than laboratory or artificial settings, flexible design revisions rather than fixed procedures, many dependent variables rather than one or two variables, and capturing social interaction rather than isolated learning (Barab & Squire 2004).

When performing design-based research, one implements one's theory and materials in a real classroom environment. This environment is spontaneous and messy. For this reason, flexible design revision may need to occur in order to assist and improve the learning environment. Because design-based research views educational interventions as holistic (Design-Based Research Collective 2003), there are many dependent variables that play a part in the outcomes. For instance, in an implementation, the interactions between the materials, activities, teacher, and students all influence learning outcomes. Finally, because the implementation involves multiple variables that work together in an active classroom, design-based research captures social interactions rather than secluded learning.

In design-based research, an impact is made on both the context being studied and on the theoretical work to which one is trying to contribute. For instance, in an inquiry-based project called the *Virtual Solar System Project*, Hay (2002) was able to make an impact on the learning environment through student learning gains and student collaborations, and he was able to contribute to cognitive psychology with claims about project-based learning and situated cognition (Barab & Squire 2004). Furthermore, with design-based research, any lessons learned can be generalized to different learning environments. Thus, efforts have both immediate and continuing consequences. The immediate refinements incorporated into the learning environment help the students within the context studied, while plans for future refinements contribute to future educational implementations as insights that can help change curriculum, materials, and teaching methods (Clarke 2003). This study examined the implementation of *Astronomy Village* and the accompanying curriculum and assessment materials and observed ways to modify the materials and learning environment to help support and improve student argumentation.

## 4. METHOD

### 4.1 Participants

The software and materials were implemented in an 11th- and 12th-grade astronomy classroom in Georgia. The classroom was made up of 15 students, 9 males and 6 females. Students were divided into five groups of three and worked together on activities and feedback conversations. Two of the five groups were videotaped while participating in the feedback conversations in order to form an understanding of how the students in the classroom were engaging in argumentation. The teacher was asked to form two groups, one group of students with low interest in astronomy and one group of students with high interest in astronomy. Interest was assessed by the teacher. These two groups were designed to be representative of the diversity that is usually seen in an elective science class such as astronomy. For instance, many students enroll in a class like astronomy because they are highly interested in and passionate about the subject, while other students enroll in the class only to avoid taking a higher-level and more difficult science class.

### 4.2 *Astronomy Village* and Authentic Learning

One way to help support students' argumentation skills is to provide them with tasks and activities that engage them in the same culture and study as actual scientists (Shin & McGee 2003). *Astronomy Village* does just that. The software places students in an authentic research context based on the Kitt Peak National Observatory in Arizona, enabling them to explore issues of current importance to astronomers. In addition to engaging the students in inquiry-based learning, the interactive lectures, hands-on labs, thought questions, article readings, and image processing activities that students complete engage them in many of the same types of activities in which astronomers participate. Furthermore, the activities allow students to see that astronomers do not have all the answers and are still discovering and developing theories about space. This emphasizes the ongoing and changing world of science, especially astronomy, in which findings, theories, and technologies are constantly being refined and improved. This is expected to encourage students to see how memorizing facts without an understanding of the ideas behind those facts is meaningless in a discipline in which ideas and current understandings are continuously changing.

### 4.3 Curriculum and Assessment Formation and Administration

The project started by identifying 13 of Georgia's Quality Core Curriculum (QCC) standards that were well-matched to the various investigations in the *Astronomy Village* software. A four-week, 20-hour curriculum was then developed that targeted the 13 standards. This resulted in 4 of the 10 *Astronomy Village* investigations being included in the curriculum. In addition, within each investigation, the 11 to 23 possible activities were examined, and a subset was identified as most relevant based on their application to the investigation topic, their ability to be completed in four 50-minute class periods, and their alignment to the standards. For each of the 13 state standards, corresponding standards from the local system and national standards were also selected.

Central to the project were three levels of assessment that were developed for this curriculum. These assessments represented a comprehensive framework designed to help align innovative curriculum with external tests by balancing formative and summative functions within and across levels. The three levels (quizzes, a curriculum-oriented exam, and a standards-oriented test) are increasingly *distal* (relative to the

specific instructional event), increasingly *summative* (versus formative), and increasingly *formal* (in terms of the representation of science knowledge).

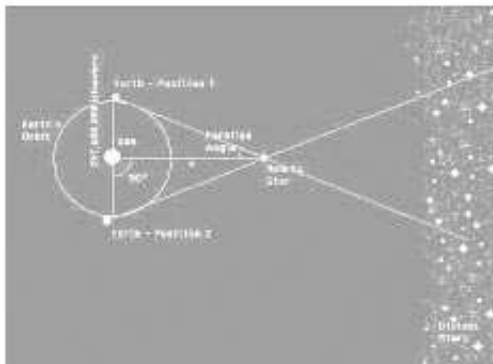
### 4.3.1 Activity-Oriented Quizzes

One quiz was developed for each of the four investigations (with each investigation lasting four days), and each quiz was administered to students the day after each investigation. The quizzes were "activity-oriented" because they were closely aligned to the specific activities that students had completed. The quizzes assessed the main concepts of the week's investigation while covering the QCC standards studied that week. The items on the quizzes addressed content covered only in *Astronomy Village* and used screen shots, wording, and diagrams directly from the software (see Figure 1 for a quiz item from week two's investigation, "Search for Nearby Stars"). Finally, because the quizzes focused on concepts and content targeted only in *Astronomy Village* and used similar (if not identical) wording and images, the quizzes were a near transfer of knowledge from the curriculum.

#### QUIZ ITEM

**2a** An astronomer is gazing at two stars in the night sky. He wants to see which of the two stars is closer to the Earth, so he measures their parallax. He finds that star 1 has a parallax of 0.5 seconds of arc, while star 2 has a parallax of 0.2 seconds of arc. Which of the two stars is closer to the Earth?

**2b** Explain your choice.



**Figure 1.** Sample Quiz Item from Week Two.

Each two-part item on the quiz required students to make claims and then warrant those claims using data. Upon the completion of each quiz, students assembled into groups to discuss the ungraded items. The quizzes were largely formative and were not intended to be graded; rather, they were intended to provide feedback to guide the enactment of a "feedback conversation" based on the activities that students had just completed. These feedback conversations were structured around a four-step argumentation review routine

that was designed to help structure and facilitate student argumentation (see Figure 2). These instructional steps were accompanied by answer explanation rubrics. Students were given the argumentation review steps, and for each item, students were told to first state their individual answers to their group members and then warrant or explain their answers using data. The group would then try to come to an agreement on the most "sensible solution."

### FOR EACH QUIZ AND EXAM ITEM

- 1. EXPLAIN AND COMPARE EACH ANSWER:**  
Everyone should *share* what they wrote earlier and *explain why* they wrote it or how they knew.
- 2. REACH INITIAL CONSENSUS:**  
The group should try to agree on the most sensible solution or agree to disagree. Everyone should understand why one or more answers are sensible.
- 3. REVIEW ANSWER EXPLANATION:**  
As a group, read aloud the answer explanation sheet. Use the written explanation and the language of science to agree on the most sensible solution.
- 4. CONFIRM GROUP UNDERSTANDING:**  
Everyone should understand and agree with a final sensible solution. Take your time. It's more important that everyone understands than to finish quickly. Also, consider why the item was written in the first place.

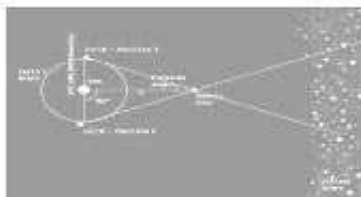
**Figure 2.** Argumentation Review Steps.

Next, students were provided answer explanations (see Figure 3 for an answer explanation for week two) that detailed the rationale underlying item solutions without explicitly stating the answers. Students were told to read them aloud and consider them relative to their consensus understanding. Finally, in order to agree on and understand a final, most sensible solution, each group was instructed to use the answer explanations to form the understandings needed to further warrant their claims using data.

Although astronomers cannot rely on brightness to measure distances to stars, they can use the concept of parallax to determine how distant a star is from the Earth. An astronomer can look at two different stars in the sky and see that one shows a great degree of parallax, while another shows a very small amount or even an immeasurable amount of parallax. From the lecture, we learned that parallax depends on distance, with nearer objects showing greater parallax than more distant objects. Furthermore, very distant objects show no parallax. Therefore, stars with a greater degree of parallax are closer to the Earth than stars with a much smaller degree or no degree of parallax.

To help you get a better understanding of this, think back to the article "Distances to Nearby Stars." Hold your finger a few inches from your face and look past it at distant objects. When you quickly close one eye and then the other, your finger appears to change position, or "shift," when compared with the objects in the background, because each eye is viewing the object from a slightly different position. Nearby stars also experience the same shift. However, if you hold your finger farther away from your face, you will see that this shift is much smaller than when you held your finger close to your face. So you can see how distant stars show smaller, or even no shift, and are often used as background reference points since they don't shift.

It is very important to remember that your eyes represent the Earth at two different positions as it revolves around the Sun, and your finger represents the star that seems to shift as the Earth rotates around the Sun.



Take a look at this picture and think of your left eye as Earth Position 1 and your right eye as Earth Position 2. As the Earth moves around the Sun, the nearby star seems to shift.

**Figure 3.** Answer Explanation from Week Two.



Much of standard classroom assessment is focused on assessing students' memorized "one right answer." However, for the open-ended quizzes, students were told to discuss their responses and agree on the most sensible solution. The purpose was to encourage students to explore and discuss the possibilities of different explanations rather than focus on coming up with a single right answer. On the quizzes, the students could suggest various claims as long as they were able to warrant those claims. Thus, it was expected that the argumentation review steps and the format of the quizzes and answer explanations would help students engage in quality argumentation as they worked to explain and compare their responses, argue and defend their own ideas, analyze each other's ideas, review the answer explanations, and use the data from the answer explanations to form a consensus and understanding of the most current and viable explanations.

### ***4.3.2 Curriculum-Oriented Exam***

The exam was "curriculum-oriented" because it was designed to assess the broader content that was targeted by the activities and reviewed in the quizzes. The exam was aligned to 13 QCCs targeted by the curriculum, and one question per standard was cherry-picked from various public sources and specifically chosen to address concepts that would be covered during the four-week implementation. The exam was administered to students twice: before the curriculum as a component of the pretest (standards-oriented test) and upon the completion of the curriculum. This allowed for a highly sensitive measure of the students' learning and retention at the end of the curriculum. Furthermore, although the questions for the exam were chosen to address concepts that would be covered during the four-week implementation, because they were selected from various public sources, the exam functioned as a pre-post measure of medium transfer learning gains.

The students completed the postexam immediately following the feedback conversation after the fourth and final quiz. Because the postexam was designed to be used for assigning student grades for the unit and to guide formal revision of the curriculum and formal remediation for specific topics and for specific students, the exams were graded and returned to the students the next day. Upon receiving their exams, students spent the day engaged in argumentation over their exams.

### ***4.3.3 Standards-Oriented Test***

The test was "standards-oriented" because it was aligned to the targeted science standards rather than the particular curriculum. The test was assembled in a manner that would allow for valid comparisons across any curriculum that targeted the specified standards, permitting valid estimations of the impact on the assessed students' performance on any high-stakes test items intended to assess those same standards. In developing the test, an astronomy content item pool of 101 multiple-choice questions was created by aligning 2 to 30 astronomy questions from various public sources to the 13 QCC standards. From this pool, two questions for each of the 13 QCCs were randomly selected, forming a 26-item test.

In addition to the 26 astronomy questions on the test (four of which were inquiry-based), the test included three "inquiry triangle" questions drawn from a pool of inquiry-related items. The inquiry triangle was made up of three inquiry items that were not tied to any particular subject area. The inquiry triangle was included to examine the influence of *Astronomy Village's* inquiry-based learning environment in improving students' ability to solve astronomy and general inquiry-type problems. In all, the test consisted of a total of 29 questions, and because the item pools were designed with a focus on the standards and not the curriculum, the test included concepts not necessarily covered in the *Astronomy Village* curriculum.

The test was formally administered to the students by a research assistant four days before the curriculum was implemented and was administered by the same research assistant four days after the completion of the curriculum. The test was administered under such secure conditions to allow us to make valid claims about gains on high-stakes achievement tests. The only information that the teacher received about the test was the students' overall test scores, as is the case with standardized assessments. Finally, because the test items were randomly selected and included items not necessarily covered in the *Astronomy Village* curriculum, the test functioned as a pre-post measure of far transfer learning gains.

#### **4.4 Assessment Design**

Having three levels of assessment made it possible to balance the formative and summative functions of assessment. Although the quizzes and exam allowed for argumentation and student feedback so as to immediately and directly benefit the students, the standards-oriented test emphasized and reflected the summative function of high-stakes assessments by focusing on state standards. Thus, the three levels of assessment allowed the opportunity to bridge the gap between the "immediacies" of a technology-based learning environment and the "abstraction" of standards (Zuiker & Hickey 2004). The goal was to engage students in argumentation over the quizzes and exam to help improve their understanding of how to engage in authentic science, which was expected to "promote a more refined understanding of the goals of the activity" (Felton & Kuhn 2001, 137) and to ultimately raise scores on high-stakes assessments focusing on state standards (Zuiker & Hickey).

### **5. RESULTS**

#### **5.1 Coding of Tapes for Argumentation**

To form an understanding of the quality of the argumentation in the classroom over the four-week period and how the students' argumentation influenced learning gains, two of the five groups of students—one group of high-interest students and one group of low-interest students—were selected, and their discussions over the four activity-oriented quizzes were videotaped and analyzed. Although the groups were formed based on interest, the students' interest strongly corresponded to their previous semester astronomy grades. The three low-interest students fell among the lowest grades in the class. Of the three high-interest students, one received the highest grade in the class, one received the third highest grade in the class, and the third student received among one of the lowest grades in the class. For the low-interest students, performance and interest were closely related, but this was not necessarily the case for the high-interest students.

The goal was to determine the makeup of the discussions of the two groups and determine if there were any differences in the quality of argumentation of these two disparate groups and if there were any improvements in the quality of the arguments with the passing of each week. The transcripts were coded in accordance with Toulmin's (1958) model. Therefore, the arguments were transcribed and analyzed in terms of students making claims, using data to support their claims, using warrants to provide a link between the data and the claim, using backings to strengthen the warrants, and using rebuttals to point to the circumstances under which the claim would not hold true. This was done by examining the transcripts and categorizing the students' statements in terms of how well they fit into each category. Conversational turns, or how the students' statements fit into the context of the discussion, were considered in order to categorize the statements more accurately. Examples from the students' discussions include:

- Claim: "The star will blow up one day."
- Data: "At each peak [on the chart], it's the same place it was six days before."
- Warrants: "Star one is closer because you cannot cover it all."
- Backings: "UV rays are harmful so they must have higher energy."
- Rebuttals: "Not necessarily."

Transcripts were read independently by two coders who then met to compare their analysis and resolve differences in interpretation. The focus was on the argumentation rather than the content of the argumentation because it was expected that engaging students in the correct process of argumentation would eventually result in them learning the science content and skills (American Association for the Advancement of Science 1993).

Compared with the low-interest group (Table 1), the quality of the argumentation of the high-interest group was superior (Table 2). The high-interest students used more data to support their claims, were much more likely to warrant their claims using data, and used more backings and rebuttals than the low-interest students. On the other hand, the low-interest students' use of claims, data, and warrants remained relatively stable throughout the four weeks, while for the high-interest students, their use of claims increased with the passing of each week. By the fourth week, the number of claims that the group was using was very high. This increase in the number of claims meant that students were just stating claims rather than warranting those claims using data.

**Table 1.** Low Interest Students' ( $n = 3$ ) Arguments over the Four-Week Period.

	Week 1	Week 2	Week 3	Week 4
Claims	8	7	6	6
Data	2	4	1	4
Warrants	1	2	1	1
Backings	0	0	0	0
Rebuttals	0	0	2	2

**Table 2.** High-Interest Students' ( $n = 3$ ) Arguments over the Four-Week Period.

	Week 1	Week 2	Week 3	Week 4
Claims	6	7	10	16
Data	3	4	6	4
Warrants	3	7	8	2
Backings	1	0	2	0
Rebuttals	0	0	8	7

One major improvement occurred in the quality of the students' arguments in that during the third and fourth weeks, both groups of students began to incorporate rebuttals into their arguments. Simon et al. (2002) emphasized that rebuttals are essential to the quality of an argument because they help change the ideas and thinking of the other individual. Conversations with rebuttals demonstrate a higher level of argumentation because they change ideas. However, aside from the increase in the use of rebuttals (and backings for the high-interest students), the quality of the conversations did not improve over the four-week period.

## 5.2 Classroom Learning Gains

Data were collected from the pre-postexam and pre-posttest for the entire classroom. Two items on the preexam resulted in a ceiling effect, so only 11 items were analyzed for the pre-post exam. Of the 11 items that were the same on the preexam and postexam, the 12 out of 15 students who took both administrations of the exam received an average score of 7.17 out of 11 on the preexam and 9.42 out of 11 on the postexam. Therefore, the 12 students' exam score increased by an average of 2.25 questions (Table 3). Of the 29 items on the pre-post test, the 14 out of 15 students who took both administrations of the test received an average score of 15.29 out of 29 on the pretest and 16.14 out of 29 on the posttest. Therefore, the 14 students' test scores increased by an average of .85 questions (Table 3).

**Table 3.** Students' Exam and Test Scores.

	Pre Mean (SD)	Post Mean (SD)	Average Gain
Curriculum-oriented Exam ( $n=12$ )	7.17 (2.04)	9.42 (1.78)	2.25
Standards-oriented Test ( $n=14$ )	15.29 (3.36)	16.14 (2.66)	.85

Statistical analyses indicated that the exam gains,  $F(1,11) = 18.43$ ,  $p < 0.001$ , were statistically significant, and test gains,  $F(1,13) = 3.21$ ,  $p < 0.097$ , were not statistically significant. Although learning gains did occur, they were very slight for the test. This may have been a direct result of the problems with the quality of the students' arguments. Thus, based on the learning gains on the test and analysis of the

argumentation of the two groups, it was clear that there was room to improve the quality of the students' arguments.

## **6. LESSONS LEARNED**

Consistent with the goals of design-based research, the implementation was examined for ways to improve the materials and learning environment to further support argumentation, and accordingly, learning gains on standards-oriented tests. Examination of the videotapes and classroom observations indicated that the duration of the discourse during the feedback conversations was very short throughout the four-week period. In addition to the duration, as can be seen in the coding of the videotapes, the quality of the discourse was also poor. Along with the videotapes of the students engaging in argumentation, classroom observations, in-depth interviews with the teacher, and self-reports from the students allowed for the opportunity to determine how to best help students engage in quality argumentation. Below is a description of the possible reasons that the students may have had problems engaging in argumentation, and potential ways to help eliminate these problems. Although two very different groups were videotaped while engaging in argumentation, it was found that the analysis of the two groups allowed for the development of ideas and materials that were useful for helping even the most extreme types of students in a classroom.

One possible factor contributing to the problem of engaging students in quality argumentation is that over the four-week period, the students illustrated uncertainty about how to use the argumentation review steps and answer explanations because they either failed to look over them or just skimmed through them quickly. This affected the structure, duration, and quality of the discussions. In fact, during the third week of the quiz feedback conversations, the quality and duration of the discourse was so low that the teacher was asked to help facilitate the students in their discussions. Although engagement time during the teacher's active facilitation increased during the two quiz feedback sessions as students increasingly engaged in the analysis of answer explanations (about two minutes more), the quality of the discussions did not increase. This was because the teacher was unsure of the best way to facilitate this new curricular routine. Much of the teacher's engagement was structured around his own routine of walking around the classroom looking for disengaged groups. He would then approach those students and ask them for their responses on the quiz. However, his way of facilitating the students who were having trouble with the quiz items did not appear to encourage worthwhile argumentation. For instance, when the low-interest group was not discussing, the teacher approached them and asked them for their responses. If the students were confused about an item, most of the scaffolding that occurred involved the teacher asking the students known-answer questions to which the students responded with the appropriate answer (given his implicit direction). This facilitation was centered on getting students to a right answer or claim rather than encouraging them to warrant the claims that he was leading them toward. For the high-interest group, the teacher approached the students if they were not engaged in discussion and asked them for their responses. If the students provided inaccurate claims or needed assistance, he would encourage them to keep talking about it.

The teacher's different approach toward the two groups may be due to his identification of the high-interest group as being more interested and thus more likely to keep discussing the astronomy. He may have felt that the low-interest students were not as likely to engage in discussion and that it would be more useful to just lead them to an answer. However, this type of facilitation appeared to contribute to the low-interest students' lack of engagement in discussion and to their problems engaging in effective argumentation. Overall, the teacher facilitated the other groups in the classroom similar to the way that he

facilitated the high-interest students; he encouraged the students to keep talking when they were having trouble answering the questions or when they were providing inaccurate claims.

Although the teacher facilitated the students differently, especially the low-interest and high-interest students, he did not examine the quality of the students' discussions by confirming that they were using the argumentation review steps and answer explanations to structure and facilitate their conversations, or that they were warranting their claims using data and supporting their claims using backing or rebuttals. This is not a surprise; again, one of the greatest difficulties in engaging students in quality argumentation is the teacher's lack of preparation for or proficiency in supporting argumentation in their classrooms. However, the teacher's role in engaging students in argumentation is critical. In fact, in their own analysis of argumentative discourse, Simon et al.'s (2002) data suggest that the amount of argumentative discourse that students engage in is more a result of the teacher's structuring and organization of the lesson than any feature or characteristic of the students or groups.

Based on classroom observations and a thorough examination of the literature on argumentation (Bell & Linn 2000; Driver et al. 2000; Duschl & Gitomer 1997; Herrenkohl et al. 1999; Kuhn 1991, 1993; Simon et al. 2002), ideas for a set of teacher facilitation guidelines were developed to help teachers engage students in scientific argumentation. These guidelines will be useful for facilitating teaching every type of student in the classroom. The teacher facilitation guidelines should encourage teachers to first listen to and observe students' discussions for three dimensions: (1) argumentation—whether students are participating in quality scientific argumentation (using data and backings to warrant claims, using rebuttals to refine claims) and are using the argumentation review steps and answer explanations to support their discussions; (2) engagement—how long the students maintain engagement in the discussion; and (3) turn taking—whether each student in the group is listening to and contributing to the discussion. Teachers should then be encouraged to approach groups or students needing assistance or not engaging in efficient argumentation with *reflective tosses*, in which the teacher hears the student or group making a claim but encourages them to warrant that claim using data; *open-ended questions* about the students' claims and warrants so that students are encouraged to further discuss and reflect on their ideas; and *content free prompts* (conjunctions: but, so) that provide useful hints that lead completely confused students to come up with the correct warrants. Therefore, the guidelines should try to move teachers away from asking, "What do I give students to develop an appropriate understanding?" to "How can I help students construct an appropriate understanding?" (Driver et al. 1994; Duschl & Gitomer 1997) so that teachers are no longer inserting content into the conversations but instead are helping students to construct understandings through the practice of argumentation.

In addition to the development of the teacher facilitation guidelines, the argumentation review steps (Figure 2) for the students should be revised to help incorporate the terminology of argumentation. For instance, rather than asking students to share what they wrote, students should be encouraged to provide claims. Rather than asking students to explain why they wrote their responses, students should be asked to warrant their claims using data. Incorporating the terminology of argumentation can help students use those terms to describe what they are doing in a way that can help them form a better understanding of the practice of argumentation.

Another concern was that although both the teacher and students would be provided guidance (teacher facilitation guidelines for the teacher and review steps for the students) on how a scientific discussion should play out, it would most likely still be difficult for them to picture what an effective feedback conversation should actually look like. For this reason, an animated video rubric is being developed to be

viewed by the students after the completion of their first quiz and feedback conversation. The video rubric will include an example of a poor conversation and an example of a good conversation based on the three dimensions of argumentation, engagement, and turn taking. In a prior study of feedback conversations in science education, it was found that providing students with video-based examples increased the proportion of conversational turns during feedback that were coded as scientific argumentation from 25% to 40% (Schafer et al. 2003). After students watch the videos, they will be asked, first on their own and then in their groups, to rate their group discussions based on the same three dimensions (argumentation, engagement, and turn taking) found in both the teacher facilitation guidelines and video rubric. The goal is to help students and teachers understand and compare good and bad scientific argumentation, to help students assess their own conversations, and to ultimately improve the duration and quality of students' argumentation.

When the implementation teacher was interviewed and asked how he thought student argumentation could be improved, he suggested adding more questions to the quiz and making the quizzes and answer explanations more difficult. He felt that making the quizzes and answer explanations longer and more complex would help engage students in longer and more involved argumentation. McCaslin & Hickey (2001) described how Vygotsky would change a task to "increase its frustrating potential, thus requiring self-directive speech" (239). Similarly, making the quizzes more challenging than students are accustomed to may increase group argumentation as students further discuss and question their ideas.

Although the students worked together in groups of three on both the activities and feedback conversations, to help students work more effectively, the teacher suggested putting students in smaller groups for the activities and larger groups for the feedback conversations. He felt that groups of two instead of three would enable students to work more comfortably and efficiently on the computer and on different activities. Meanwhile, for the feedback conversations, he suggested combining two groups, forming groups of four for the assessment feedback conversations, so that students could be exposed to a diversity of ideas.

Finally, because the students were able to complete activities more quickly than was expected, the teacher suggested adding activities from each investigation to the curriculum. Thus, in the future, it would be useful to include extra activities, particularly more interactive lectures and article readings, that provide students with further background knowledge. This additional prior knowledge gained through such activities can help the students engage in more efficient argumentation by giving them the knowledge they need to warrant their claims using data and to support their claims using backings and rebuttals.

Engaging students in argumentation is much more difficult than it appears. However, because argumentation is essential for understanding the social workings of science (Driver et al. 2000; Duschl 2000) and for learning the skills and content detailed in science education standards (American Association for the Advancement of Science 1993), it is important to develop ideas and materials that can help support scientific argumentation.

In this study, the implementation of an inquiry-based software program and supplemental curriculum and assessment materials was examined to understand how the goal to support argumentative discourse influenced the quality of the students' argumentation and their learning gains. It was found that despite the efforts to design an implementation that supports argumentation, many of the difficulties of engaging students in argumentative discourse that were found in the literature were experienced. Therefore, while trying to develop and further refine learning environment and materials to help support argumentation, it

became possible to understand and tackle many new, already-recognized challenges that present themselves when trying to engage students in argumentative discourse.

## NOTE

**Note 1.** Multimedia Educational Software (<http://www.cotf.edu>).

## 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, 1099.

American Association for the Advancement of Science. 1993, *Benchmarks for Science Literacy: Project 2061*, New York: Oxford.

Barab, S., & Squire, K. 2004, Design-Based Research: Putting a Stake in the Ground, *Journal of the Learning Sciences*, 13, 1.

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, 797.

Brown, A., & Campione, J. C. 1998, Designing a Community of Young Learners: Theoretical and Practical Lessons, in *How Students Learn: Reforming Schools through Learner-Centered Education*, N. M. Lambert & B. L. McCombs (Editors), Washington: American Psychological Association, 153.

Clarke, M. A. 2003, *Common Ground and Contested Territory in School/University/Community Collaboration: MLAR as Solution*, the Learning Performance and Support Laboratory, <http://lpsl.coe.uga.edu/>.

Cobern, W. W. 1993, Constructivism, *Journal of Educational and Psychological Consultation*, 4, 105.

Design-Based Research Collective. 2003, Design-Based Research: An Emerging Paradigm for Educational Inquiry, *Educational Researcher*, 32, 5.

Diehl, C. L. 2000, "Reasoner's Workbench" Program Supports Students' Individual and Collaborative Argumentation, Paper presented at the National Association for Research in Science Teaching, New Orleans, LA.

Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. 1994, Constructing Scientific Knowledge in the Classroom, *Educational Researcher*, 23, 5.

Driver, R., Newton, P., & Osborne, J. 2000, Establishing the Norms of Scientific Argumentation in Classrooms, *Science Education*, 84, 287.

Duschl, R. 2000, Making the Nature of Science Explicit, in *Improving Science Education*, R. Millar, J. Osborne, & J. Leach (Editors), Philadelphia: Open University Press, 185.



- Felton, M., & Kuhn, D. 2001, The Development of Argumentative Discourse Skill, *Discourse Processes*, 32, 135.
- Hay, K. 2002, *Virtual Solar System Project*, <http://inkido.indiana.edu/research/a100.html>. Link to Hay's work at the University of Georgia: <http://lpsl.coe.uga.edu/projects/ele.asp>.
- Herrenkohl, L. R., Palinscar, A. S., DeWater, L. S., & Kawasaki, K. 1999, Developing Scientific Communities in Classrooms: A Sociocognitive Approach, *Journal of the Learning Sciences*, 8, 451.
- Jimenez-Aleixandre, M. P., Bugallo-Rodriguez, A., & Duschl, R. A. 1997, *Argument in High School Genetics*, Paper presented at the National Association for Research in Science Teaching, Chicago, IL.
- Koslowski, B. 1996, *Theory and Evidence: The Development of Scientific Reasoning*, Cambridge, MA: MIT Press.
- Kuhn, D. 1991, *The Skills of Argument*, Cambridge: Cambridge University Press.
- Kuhn, D. 1993, Science as Argument: Implications for Teaching and Learning Scientific Thinking, *Science Education*, 77, 319.
- Lemke, J. L. 1990, *Talking Science: Language, Learning and Values*, Norwood, NJ: Ablex Publishing.
- McCaslin, M., & Hickey, D. T. 2001, Self-regulated Learning and Academic Achievement: A Vygotskian View, in *Self-Regulated Learning and Academic Achievement: Theory, Research, and Practice*, 2nd. ed., B. Zimmerman & D. Schunk (Editors), Mahwah, NJ: Erlbaum, 227.
- McClelland, G. 1983, Discussion in Science Lessons, *School Science Review*, 65, 129.
- Newton, P., Driver, R., & Osborne, J. 1999, The Place of Argumentation in the Pedagogy of School Science, *International Journal of Science Education*, 21, 553.
- Norris, S. 1997, Intellectual Independence for Nonscientists and Other Content: Transcendent Goals of Science Education, *Science Education*, 81, 239.
- Rudolph, J. L., & Stewart, J. 1998, Evolution and the Nature of Science: On the Historical Discord and Its Implications for Education, *Journal of Research in Science Teaching*, 35, 1069.
- Savery, J. R., & Duffy, T. M. 1996, Problem-Based Learning: An Instructional Model and its Constructivist Framework, in *Constructivist Learning Environments: Case Studies in Instructional Design*, B. G. Wilson (Editor), Englewood Cliffs, NJ: Educational Technology Publications, 135.
- Schafer, N. J., Hickey, D. T., Zuiker, S., Kruger, A. C., & Russell, H. A. 2003, *Using Video Feedback to Facilitate Classroom Assessment Conversation*, Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Shin, N., & McGee S. 2003, *Virtual Design Center*, <http://vdc.cet.edu/main.html>, Wheeling Jesuit University.

Simon, S., Erduran, S., & Osborne, J. 2002, *Enhancing the Quality of Argumentation in School Science*, Paper presented at the Paper presented at the annual meeting of the National Association for Research in Science Teaching, New Orleans, LA.

Slavin, R. E. 1995, *Cooperative Learning: Theory, Research, and Practice*, 2nd ed. Boston: Allyn & Bacon.

Sutton, C. 1992, *Words, Science and Learning*, Milton Keynes, UK: Open University Press.

Toulmin, S. E. 1958, *The Uses of Argument*, New York: Cambridge University Press, 2003.

Zhu, E. 1998, Learning and Mentoring: Electronic Discussion in a Distance Learning Course, in *Electronic Collaborators: Learner-Centred Technologies for Literacy, Apprenticeship, and Discourse*, C. J. Bonk & K. S. King (Editors), Mahwah, NJ: Erlbaum, 233.

Zuiker, S. J., & Hickey, D. T. 2005, *Assessment Technologies for Technology-Supported Science Education: Engineering Discursive Feedback in a 'Classroom of the Future'*, Paper presented at the American Educational Research Association Meeting, Montreal, Canada.

ÆR

53 - 70