Adventures in Argument: Training in Argumentation Influences Student Resource Use in Collaborative Meaning Making

Julia Gressick, Indiana University South Bend, 1700 Mishawaka Avenue,
South Bend, IN 46634, jgressic@iusb.edu
Sharon J. Derry, University of North Carolina-Chapel Hill, CB# 3500, Peabody Hall,
Chapel Hill, NC 27599-3500, derry@unc.edu

Abstract: Argumentation is the primary pedagogical strategy employed in the online undergraduate course Human Abilities and Learning Online (HAL Online). We conducted a controlled *in vivo* experiment in this course to examine the effects, on collaborative meaning making, of providing direct training in argumentation early in the course. The performance of a group receiving the treatment, *Trained Argumentation with Modest Scaffolding* (TAMS), was compared with an ecological control group that did not receive argument training: *Emergent Argumentation with Modest Scaffolding* (EAMS). We hypothesized that argument training would influence how students attended to, used, and shared instructional resources as evidence to support explanations in collaborative meaning making. Results indicated that TAMS exerted strong influence on how deeply and thoroughly students processed, were accountable to, and integrated instructional resources.

Introduction

Argumentation as pedagogical practice is widely advocated for its potential to improve learners' conceptual knowledge and ability to reason in the domains of science, mathematics and social science (e.g., Cavagnetto, 2010; Kuhn, 2010; Noroozi, Weinberger, Biemans, Mulder, & Chizari, 2012). Although there is consensus concerning the value of instructional argumentation, the literature on this topic reveals a complex landscape of perspectives on how to conceptualize and design such instruction. In his review of argument in science education, Cavagnetto (2010) identifies three forms of argument as pedagogy: direct instruction in argument structure prior to engaging in scientific activity; developing argument skills through mentoring during immersion in science process; and instructing through ethical or political dilemmas, engaging argument processes that individuals experience socially from a young age (Hay & Ross, 1982). In agreement with Sandoval (2005) and others, Cavagnetto makes a case for the immersive approach. Like Kuhn (2010), however, he recognizes that more research is needed to understand this issue. Like others (e.g., Larson, Britt, & Kurby, 2009; Schworm & Renkl, 2007), we will make a case for direct, explicit instruction in argument, at least for contexts similar to ours.

The setting for the work reported here is Human Abilities and Learning, an upper-level undergraduate course offered at a large university. It is required for many majors, including teacher education. The course addresses the scientific basis of thinking and learning and what this implies for guiding children and adults, for personal development, and for building environments that help people learn and grow successfully. Typically the course is offered as a large lecture course. However, for many semesters one professor has offered a non-traditional section (HAL Online) for students preferring a problem-based format that may meet face-to-face occasionally during the semester but is taught mostly online. The course aims to develop scientific literacy through reading and online argumentation around real-world problem tasks, often presented with video cases. Students are assigned to small groups of 3-5 members that work online throughout the semester.

The units in the spring 2011 HAL Online offering that was the data source for this study were: I. Cognition and Culture; II. The Amazing Learning Brain; and III. Using Learning Science in Reflective Practice. Each unit comprised four or five weeklong lessons. During a lesson, students study multimedia resources about psychological science content drawn from textbooks and from video and news sources such as TED.com and The New York Times. In alternate weeks students either post a reflective personal blog that answers a problem-solving prompt, or participate in online collaborative problem-solving tasks. Blog posts and discussions are graded using a rubric that rewards understanding and intelligent use of course ideas.

Since 2007, the course has been a site for field testing Video Mosaic (Videomosaic.org), an NSF-funded research and development project that has created an online repository comprising an extensive, searchable, annotated collection of research video on children's mathematical reasoning and development. This collection, based on the work of researchers Robert B. Davis and Carolyn Maher (e.g., Maher, 2005) is a valuable resource that builds on extensive prior research including a longitudinal study following the same cohort of students through high school and beyond. Work described in this paper represents research and development with this valuable resource for teacher education.

The Problem

We hypothesized that students who argue better in our course will achieve more. Yet, despite a natural tendency to argue, students' arguments are often ill formed, lack evidence, and are incomplete (Kuhn, 2005). Moreover, individuals sometimes fail to understand what qualifies as evidence (Glassner & Schwarz, 2005). Struggles with constructing sound arguments are complicated by resource-rich environments resulting from advances in technology. Integrating evidence from a variety of multi-media resources has become an increasingly important component of arguing well. Our challenge is to design instruction that promotes students' development of sound arguments in resource-rich environments.

The approach favored by science educators (e.g., Cavagnetto, 2010) is immersion with scaffolding of argument. Yet this approach has practical limitations for our online university setting where there are pressures to increase enrollments despite lower instructional budgets. Scaffolding must be provided online to many students by a single faculty member unassisted, or with the help of relatively inexperienced teaching assistants who themselves may have poor argument skills. Recently universities have begun to offer relatively unsupervised massive open online courses (MOOCs) aimed at large-scale participation and open access via the WWW. Our aspiration is to serve large enrollments using an argument-based pedagogy. It is important, therefore, to investigate the viability of pragmatic alternatives to human guidance of argument online. One option is formally training students in argument *prior* to engaging them in pedagogies that require integration of conceptual content and evidence from multiple sources as a basis for well-reasoned claims.

Toward that end we conducted a controlled experiment within HAL Online to examine the effects, on individual learning and collaborative meaning-making, of a week-long unit offered early in the course that provided direct training intended to improve students' understanding of and ability to engage in good argument. The performance of a group receiving the treatment, *Trained Argumentation with Modest Scaffolding* (TAMS), was compared with an ecological control group that did not receive argument training: *Emergent Argumentation with Modest Scaffolding* (EAMS). Except for the treatment manipulation, both groups participated in an identical course of study and assessment. One hypothesis was that achievement for *individual* students, as measured by tests of comprehension and scientific literacy, would be higher for students who participated in TAMS. This hypothesis was strongly supported and is detailed in another analysis reported elsewhere (Gressick & Derry, 2013).

In contrast, the focus of analyses reported in this paper was on whether training in argumentation can influence aspects of *collaborative* meaning-making on tasks requiring small groups to integrate scientific ideas from course material with observations from real-world situations, to create evidential arguments for scientific explanations. Our analysis addressed the following research questions: Are there differences in how students in TAMS vs. EAMS use the scientific course material in their reasoning? Are there differences in how and how closely they attend to details of real-world cases provided by the problem? Is there evidence that argument training produces differences in how members of groups work together to blend their ideas and reasoning?

Theoretical Framework

The TAMS treatment was largely implemented thorough the lesson, *Adventures in Argument*, inspired by the Toulmin (1958) model that has served as the basis for many educational approaches using argumentation (e.g. Kuhn, 2005; Means & Voss, 1996 Stegmann, Weinberger, & Fischer, 2007). Toulmin's model is the basis of Halpern's (2002) *Analyzing Arguments*, a chapter in her award-winning text *Thought and Knowledge*, which was an assigned reading constituting a portion of the TAMS treatment. Her chapter emphasizes recognizing and using five components of good argument: conclusions, premises, counterarguments, qualifiers, and assumptions.

Arguing to support scientific explanations or theories is a social practice involving communication and persuasion. Berland and Reiser (2009) draw an epistemological distinction between the process of defending scientific explanations and the process of creating them, two key but distinct components of scientific practice (Kuhn, 2010). Although constructing explanations is central to scientific practice, it is not the only goal. Emphasis on constructing explanations can even undermine attention to evidence. Training students in argumentation, we hypothesized, would direct their attention to using course material as sources of evidence to justify and support explanations.

Data Source and Design

We conducted an *in vivo* experiment, which manipulates elements of instruction in a natural setting and observes the effects on student learning (e.g. Aleven & Koedinger 2002). The context of the study was the spring, 2011 offering of HAL Online. Forty-four students enrolled. A summary of the experimental design is shown as Figure 1. The treatment manipulation was the last lesson in the first course unit. The context for studying treatment effects on collaborative meaning-making were forum discussions in units occurring four weeks and eight weeks following the treatment. A separate Moodle course environment was created for each condition. These were identical except for the treatment-related manipulations.

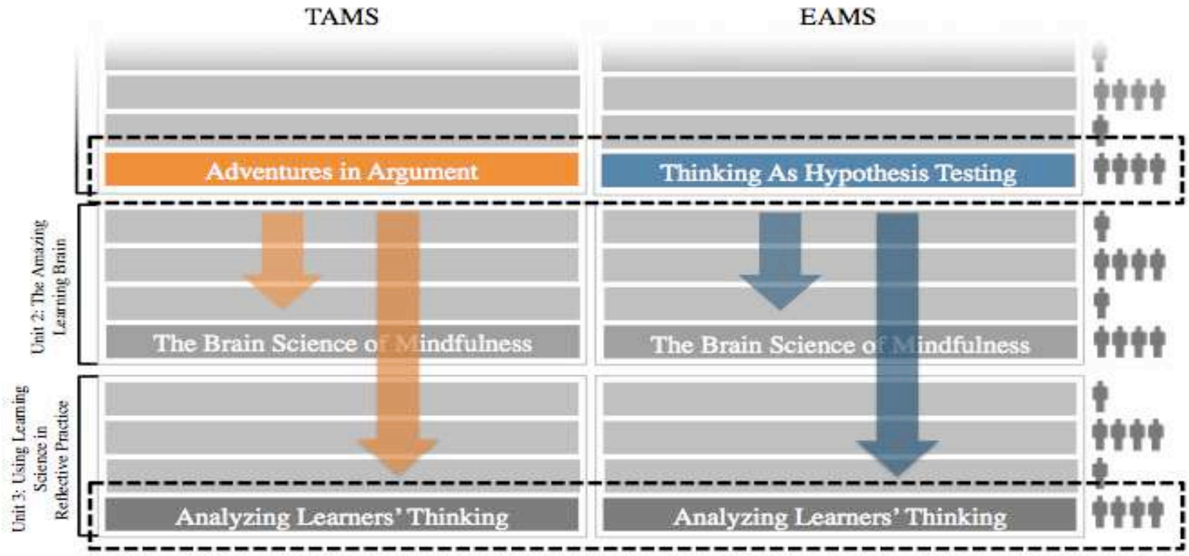


Figure 1. Overview of instructional design for TAMS and EAMS.

Using a within classroom nested design (Salden & Koedinger, 2009), students enrolled in HAL Online were assigned to small groups based on common interests as determined by self-report surveys. Small groups were randomly assigned to two conditions (described next). Groups comprised three or four students and, to avoid confounding the group dynamic, were maintained throughout the semester.

Treatment: Training in Argumentation with Modest Scaffolding (TAMS)

The goal of this lesson was to teach students the skill of making and recognizing strong arguments. Students read "Analyzing Arguments," a 50-page chapter on argumentation from *Thought and Knowledge* (Halpern, 2002). Following a quiz, students engaged in a collaborative forum discussion with their small group in which they practiced using ideas from reading to support analyzing and evaluating an argument in a speech.

Ecological control: Emergent Argumentation with Modest Scaffolding (EAMS)

In the EAMS control group students received an alternative weeklong lesson that focused on an alternate chapter of similar length and density from *Thought and Knowledge*, "Thinking as Hypothesis Testing." This chapter presented topics like the nature of variables, correlational versus experimental evidence, and using evidence for causal claims. During EAMS students completed a quiz and participated in a collaborative forum that employed the identical video speech but required designing a study to test the speaker's causal claims.

Description of Collaborative Forum Activities

Data were analyzed from two online collaborative forum activities during the semester. The first was titled *The Brain Science of Mindfulness.* This lesson occurred at the end of unit that directly followed the experimental argumentation (or ecological control) lesson (See Figure 1). Students read and viewed video about the scientific study of meditation practice. In their forum they debated the scientific merits of a proposed meditation-training program for a struggling middle school and were required to reach a group consensus.

The second forum analyzed, the primary focus of this paper, was in the lesson *Analyzing Learners' Thinking for Evidence of Preparation For Future Learning* and was a significantly more complex task that required sophisticated integration of multiple text and video sources. This lesson occurred as the final collaborative forum at the end of the course (see Figure 1). The primary goal was for HAL Online students to bring their knowledge of the claims of constructivist education together with claims about the nature of transfer proposed in a theory by Bransford and Schwartz (1999) to help them collaboratively examine elementary students' problem solving over time and in depth for the purpose of evaluating the scientific claims of the theory. This assignment represented a case of Berland and Reiser's "defending" or "persuasion" component of science practice, which they distinguish from constructing explanations. The explanations they were evaluating had been developed previously based on resources already encountered in Unit 3: *Rethinking Transfer* (Bransford & Schwartz, 1999) and *Should Schools Adopt a Constructivist Approach to Education?* (Windschitl & Hirsch, 2002). To seek evidence for these theoretical explanations, students were directed to access a series of six video clips from the Video Mosaic repository (Videomosaic.org), in which 11th grade students solved and justified solutions to a combinatorics problem, leading them to struggle with understanding Pascal's triangle and exponential reasoning. The 11th graders in the videos had been part of a cohort followed from early grades and that had been immersed in constructivist learning environments through their years of schooling. HAL students had previously studied videos of this same cohort solving and justifying similar combinatorics problems within a constructivist educational setting in the 4th grade. The HAL Online students' discussion task was framed as follows and they were required to reach consensus.

> *What evidence do you find that early educational experiences have prepared the students in these video clips for future learning? How confident are you regarding claims that these students' earlier educational experiences have had an impact? What convinces you or would convince you?*

Method of Analysis

This study used analysis procedures for quantifying qualitative analyses recommended by Chi (1997) and outlined in the following stages.

First, forums were searched for instances where groups used concepts from the readings and evidence from the video series. As data were searched, a chain of reasoning (Chi, 1997) for each group was developed to eliminate multiple coding of the same concept within the same discussion thread. Because of the interconnected nature of collaborative meaning making, the group was viewed as the unit of analysis. However, connections to individual contributions are not lost in analysis and can be viewed on graphs in the results section of this paper.

Next, a sample of data was coded using the scheme in Table 1. The coding scheme follows phenomenon-based hypothesis coding (Miles & Hubermann, 1994; Saladña, 2009) that focused on the elements students leverage from course readings and the general ways that students applied these ideas to their discussions and, as a small group, engaged in meaning making. Our scheme was based on coding developed by Pena-Shaff & Nicholls (2004) and methods of collaborative meaning making described by Stahl (2006). Once the coding scheme was stabilized through discussion within our research group, reliability between two coders was calculated and reached 95% (Cohen's kappa .93) after two rounds of coding and discussion (Cohen, 1968).

Table 1. Description of coding scheme.

How do groups *integrate ideas* from text resources in discourse?		
Code	**Description**	**Example**
Quote text resource directly	Provides a direct quotation from an assigned reading	*"Information presented in the context of solving problems is more likely to be spontaneously utilized than … simple facts." (Rethinking Transfer)*
Restate text resource	Restates an idea from reading in own words	*The Bransford & Schwartz article mentions SPS testing… often fails to capture transfer…*
Extend concepts from text resources	Extends or applies a concept from reading to discussion	*The article on transfer also mentions the importance of this…for transfer to occur, the children cannot think of the concept in only one type of situation…*
How do groups *integrate ideas* from video of student learners in discourse?		
Quote video evidence directly	Provides a direct quotation from the video data	*Shelly says "my teacher's going to kill me because she knows I can do this."*
Restate video observations	Restates or summarizes observations from video	*[Robert] had already been sketching the tower problem from before this discussion was even underway.*
Extend video observations	Extends or interprets video data	*[Robert's] ideas are important to the group discussion and thinking in this clip.*
In what ways do groups *build* inter- and intra-subjective *understanding*?		
Up-take	Restates what another group member stated in prior post	*I think that the students definitely used their previous learning in these videos as well.*
Elaborate	Extends an idea previously mentioned by another group member in a previous post	*…I, too, thought about the concept "use it or lose it." In past readings about the aging brain, we read that the brain retains its plasticity…*
Draw personal connection	Makes a connection between content of discussion and personal experience	*They very quickly start with the triangle approach that we did in class and … related it to their previous experience.*

After coding, data were quantified and represented in a tabular format in order to find differences in patterns between treatment groups. Bar graphs were created to further reveal and enable study of patterns across groups and conditions. For each code we calculated group means, standard deviations, and treatment effect sizes. Our decision to focus on descriptive rather than inferential statistics was influenced by the relatively small number of groups in each condition.

Results and Discussion

Results of analyses from *The Brain Science of Mindfulness* (which did not require integration of video evidence) are briefly summarized as they replicate and strengthen findings. Our major focus is on the *Analyzing Learners' Thinking* forum. In both, TAMS clearly influenced student accountability to resources during meaning making.

Meaning-Making in the *Brain Science of Mindfulness* Forum

How Groups Used Resources

In both treatment conditions, the most consistently observed method of integrating ideas from the course was to extend the findings of research studies discussed in course readings to the case presented in the forum discussion task. The mean frequency of idea extensions for TAMS (M = 4.83, SD = 2.32) was higher than for EAMS (M = 3.0, SD = 1.67). The effect size for this analysis (d = 0.92) exceeded Cohen's (1988) guideline for large effect size (*d* = .80). Additionally, groups integrated ideas by directly quoting resources. Overall, groups in TAMS demonstrated a higher number of direct quotes in their discussions (M = 3.5, SD = 1.52) than groups in EAMS (M = 1.83, SD = 2.64), although the group with the most quotes was from EAMS. The effect size for this analysis (*d* = 0.80) was large. In TAMS, groups restated information from text resources from one to nine times over the course of the discussion, with 50% of groups demonstrating at least 5 instances of restating experts (M = 4.67, SD = 3.67). In EAMS, however, all groups produced five or fewer instances of restating experts (M = 3.17, SD = 1.17). The effect size (*d* = .62) was moderate.

How Groups Built Inter- and Intra-subjective Understanding

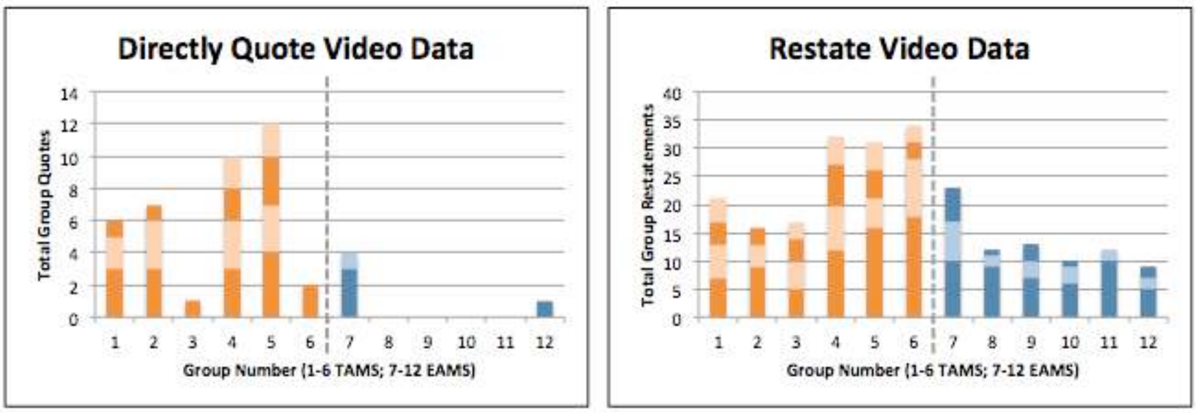
Groups in TAMS exhibited more up-take of ideas (M = 5.83, SD = 4.54) than EAMS (M = 3.17, SD = 2.23). The effect size for this analysis (*d* = .79) was high. Furthermore, groups in TAMS demonstrated more instances of elaboration of ideas (M = 2.33, SD = 2.07) than groups in EAMS (M = 0.5, SD = 0.84). The effect size for this analysis (*d* = 1.26) was large. In addition, all TAMS groups drew at least one personal connection, with most groups (67%) making at least 4 personal connections (M = 3.67, SD = 1.75). While 67% of groups in EAMS drew personal connections, only one group (Group 10) made more than one connection (M = 0.83, SD = 0.75). In most cases in TAMS but not EAMS, the personal connections made by group members were acknowledged and integrated into group discourse. The effect size for this analysis (*d* = 2.27) was large.

Meaning-Making in the *Analyzing Learners' Thinking* Forum

How Groups Used Resources

Groups in TAMS demonstrated a higher use of direct quotations from both text resources (M = 2.50, SD = 2.07) and from video data (M = 6.33, SD = 4.32) than groups in EAMS (text resources M = .1.17, SD = 1.6; video data M = .83, SD = 1.60). The effect sizes for both quotations of text resources (*d* = .72) and quotations of video data (*d* = 1.86) were large. Figure 2 provides a visual comparison of direct quotations from video data. Variations in color in each bar graph indicate contributions that were made by individual members of the group.

We examined how often groups restated (in their own words with conceptual correctness) ideas from text resources and video observations in their forum discussions. In both TAMS (M= 25.17, SD = 8.08) and EAMS (M=13.17, SD = 5.04), groups more often restated observations of the learners depicted in the videos than they did information from text resources (TAMS M=10.17, SD = 6.31; EAMS M = 4.0, SD = 1.26). Moreover, groups in TAMS for both types of resources made more restatements than in EAMS. The effect sizes for both text resource restatements (*d* = 1.63) and video restatements (*d* = 1.83) were large. Figure 3 provides a visual comparison of restatements compared across groups from video data.

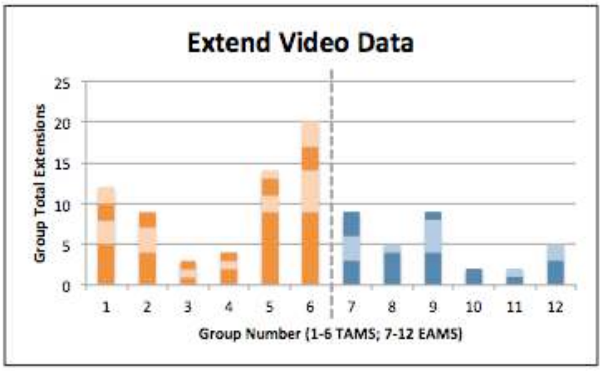


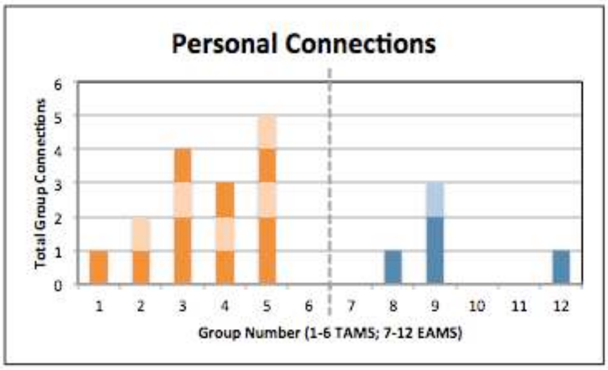
Figures 2 and 3. Direct quotations and restatement of video data resources in TAMS and EAMS.

We asked how resources were extended by groups as they constructed meaning in their discussion. Groups in TAMS were more likely to extend video (M=10.33, SD = 6.40) than EAMS (M = 5.33, SD = 3.14). Similarly, groups in TAMS (M = 7.33, SD = 4.84) were more likely to extend ideas from text resources than groups in EAMS (M = 4.0, SD = 2.0). The effect sizes for both text (d = .94) and video (d = .99) resources were large. Figure 4 provides a visual comparison of extensions of video data compared across groups.

How Groups Built Inter- and Intra-subjective Understanding

Similar to the findings for the *Brain Science of Mindfulness* forum, groups in TAMS demonstrated a higher average frequency of up-take of ideas (M = 8.50, SD = 7.42) than EAMS (M = 5.17, SD = 2.99). The effect size (*d* = .66) was moderate. Further, groups in TAMS demonstrated more instances of idea elaboration (M = 7.33, SD = 7.20) than groups in EAMS (M = 3.33, SD = 2.07). The effect size for this analysis (*d* = .88) was large. In addition, all but one TAMS groups drew at least one personal connection (M = 2.5, SD = 1.87). While 50% of groups in EAMS drew personal connections, only one group, Group 9, made more than one (M = 0.83, SD = 1.17). The effect size for this analysis (*d* = 1.10) was large (Figure 5).





Figures 4 and 5. Extensions of video data resources and personal connections in TAMS and EAMS.

An Illustrative Case of Precise Resource Use in TAMS

As presented in the data above, groups in TAMS demonstrated a more robust and precise use of resources in their discussions. The data presented in Table 2 exemplifies how precise, accountable use of data by a group's initial posting member served as a point of entry for other members of the group to participate in the data analysis and encouraged a collaborative meaning-making process. Subject 1 of Group 4 (TAMS) started the discussion with a specific set of "field notes" from his video study (see Table 2, post 2.1 below), including time stamps and direct quotes from the video clips. After this initial post, other group members may have modeled this first poster, adopting similar approaches in their own posts. In her response, Subject 43 offered an alternate, more abstracted organization of analysis focused on the participants from the video data (Table 2, post 2.3). What followed was a transformed analysis and synthesized summary by Subject 32 (Table 2, post 2.6).

Table 2. Abbreviated Group Forum Posts, TAMS Group 4

Post 2.1, Subject 1 (Initial Post)
I'm going to type my thoughts as I watch the videos to start off the discussion… ***Clip 1:*** The students are VERY aware that from somewhere, they have learned the tools to approach this type of problem. (Shelly says "my teacher's going to kill me because she knows I can do this")… ***Clip 2:*** Stephanie and Shelly are discussing how they are organizing the cases… ***Clip 3:*** I can't…understand what the teacher is asking…But at about 4:25 in this video … ***Clip 4:*** Although I'm finding it difficult to re-articulate Stephanie's explanation for why the Pascal's triangle explains the pizza situation, it's sounding pretty logical and convincing.. ***Clip 5:*** Stephanie shows an indication of transfer right off the bat: "we worked on it in 8th grade"… ***Clip 6:*** Amy gets a little bit more engaged in this one… At about 9:20 in the final clip I finally became aware of why there is a "2" as the base in the exponential expression describing the pizza problem…
Post 2.3, Subject 43: Okay so I started out taking notes on each clip like [subject 22] and [subject 1], but I came to basically the same conclusions…so..I'm…gonna add some observations about the group and each kid in it....
Post 2.6, Subject 32 (Group Summary)… drawing a connection between the Tower problem and Pascal's Triangle allowed first Robert and then the rest…to better understand the problem, both through abstract verbal reasoning and spatial imaging, and…the students use the abstract model to both…justify answers to previously solved problems and reconcile [current] confusion, which is a sophisticated and enduring form of knowledge…

In contrast, Subject 42 of Group 11 (EAMS), initially posted an abstracted, summative analysis of the data, organized around themes (see Table 3). Unlike the approach taken by Subject 1 in Group 4, Subject 42's original post relied heavily on her restatement and summary of observations from the video. This initial post lacked the accountability and precision with which Subject 1 leveraged the video resources. While there are some positive qualities of Subject 42's post, other members of the group did not engage in analysis as actively or in-depth as was observed across members of Group 4. One suggestion of why this might be is the degree to which Subject 42 had abstracted her observations to support the claims in the text. Because of this, a clear model of her meaning-making process – a point of entry into the collaborative process – was not provided to the other members of the group, as it was in Group 4. What had resulted from the approach adapted by Group 4 was a rich, integrated understanding of how specific evidence from the video could be used to promote an argument in support of the preparation for future learning theory of transfer. This example suggests the importance of precise resource use as a means to facilitate collaborative meaning-making processes.

Table 3. Abbreviated Initial Group Forum Posts, EAMS Group 11

Post 3.1, Subject 42 (Initial Post)
I think these students' prior experience with the constructivist approach…has prepared them for future learning. ***Transfer***…They used algebra that they had learned prior to this lesson to help them solve the problem…This shows that the students were able to transfer information they had previously learned…they continued to double check it and try it from other angles. This shows constructive transfer because… **Expert**… Stephanie proved that she had a deep understanding of the problem because she was able to understand and recognize almost immediately that the way the other girl had done the problem was correct even thought it was different than her way. This shows Stephanie exhibiting skills that an expert would exhibit. **Mix**…In the very beginning the students Stephanie and other girl (don't remember her name) discussed how they were going to go about solving this problem…This also shows that their knowledge was transferred… organizing thoughts and looking for deeper concepts than just a formula show skills of an expert.

Conclusions and Scholarly Significance

Discussions in groups that received argument training were consistently more enriched by references to and sharing of the course material. In the more complex problem that involved analyzing videos of children's problem solving over time to determine the credibility of a scientific theory, students trained in argument correctly incorporated into their discussions more scientific material, and they conducted more exacting and careful search of videos to identify evidence related to the theory. A formal course of training in argument might well result in more accountable discussions, where groups integrate more from resources into their discussions. That they use more direct quotes, for example, suggests attention to preserving the words of credible sources and may indicate a heightened precision regarding data from sources. The findings of this study, which demonstrate effects on student resource use occurring many weeks after *Adventures in Argument*, indicate that training prepared students for future learning (Bransford & Schwartz, 1999). The study contributes to a discussion on how to optimally approach argumentation as pedagogy and provides support for the direct training approach, at least for college-level learners in online environments that employ argumentative pedagogy. Moreover, this approach offers a viable alternative to more resource-intensive immersive approaches for online environments. While the unit of analysis in this study was the small group, data for individual students was visually supplied. The patterns of individual involvement open an area that requires further investigation: although many groups experienced participation from multiple members, there was room for improvement. However, combined with a companion study (Gressick & Derry, 2013) showing positive effects of training on individual student learning, this research shows that direct training in argumentation is a promising intervention.

References

Aleven, V., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science*, 26, 147-179.

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of Research in Education* , *24*, 61-101. Washington DC: American Educational Research Association.

Berland, L. & Reiser, B. (2009). Making sense of argumentation and explanation. *Science Educatio*n, 93, 26-55.

Cavagnetto, A.R. (2010). Argument to foster scientific literacy: A review of argument interventions in K-12 science contexts. *Review of Educational Research*, 80(3), 336-371.

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: a practical guide. *Journal of the Learning Sciences, 6*(3), 271–315.

Cohen, J. (1968). Weighed kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin* 70 (4): 213–220. doi:10.1037/h0026256. PMID 19673146.
Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (second ed.). Lawrence Erlbaum.
Glassner, A. & Schwarz, B. B. (2005). The antilogos ability to evaluate information supporting arguments. *Learning and Instruction, 15,* 353-375.
Gressick, J. & Derry, S.J. (2013). The Influence of Training in Argumentation on Students' Individual Learning Outcomes. In Rummel, N., Kapur, M., Nathan, M, & Puntambekar, S. (Eds.), *To See the World and a Grain of Sand: Learning across Levels of Space, Time, and Scale: CSCL 2013 Conference Proceedings* (Vol. 2, pp. 38 - 41). International Society of the Learning Sciences.
Halpern, D. (2002). Thought and Knowledge (4th ed.). Mahwah, NJ, Erlbaum.
Hay, D.F., & Ross, H.S. (1982). The social nature of early conflict. *Child Development*, 53, 105-113.
Kuhn, D. (2010). Teaching and learning science as argument. *Science Education*, 810-824.
Kuhn, D. (2005). *Education for thinking*. Cambridge, MA: Harvard University Press.
Larson, A.A., Britt, M. A., & Kurby, C. (2009). Improving students' evaluation of informal arguments. *Journal of Experimental Education, 77*(4), 339-365.
Maher, C.A. (2005). How students structure their investigations and learn mathematics: Insights from a longitudinal study. *Journal of Mathematical Behavior,* 24, 1-14.
Means, M. L. & Voss, J.F. (1996). Who reasons well? Two studies of informal reasoning among children of different grade, ability and knowledge levels. *Cognition and Instruction,* 14, 139-178.
Miles, M.A. & Hubermann, A.M. (1994). Qualitative data analysis (2^{nd} ed.). Thousand Oaks, CA: Sage.
Noroozi, O., Weinberger, A., Biemans, H.J.A., Mulder, M., & Chizari, M. (2012). Argumentation-Based Computer Supported Collaborative Learning (ABCSCL): A synthesis of 15 years of research. *Educational Research Review*, 7, 79-106.
Pena-Shaff, J & Nicholls, C. (2004). Analyzing student interactions and meaning construction in computer bulletin board discussions. *Computers & Education*, 43 (3).
Saladaña, J. (2009). *The Coding Manual for Qualitative Researchers*. Los Angeles: Sage Publications.
Salden, R. J. C. M. & Koedinger, K. R. (2009). In vivo experimentation on worked examples across domains. Symposium at the Thirteenth Biennial Conference of the European Association for Research on Learning and Instruction.
Sandoval, W. (2005). Understanding students' practical epistemologies and their influence. *Science Education*, 89, 634-656.
Schworm, S. & Renkl, A. (2007). Learning argumentation skills through the use of prompts for self-explaining examples. *Journal of Educational Psychology*, 99, 285-296.
Stahl, G. (2006). *Group Cognition: Computer Support for Building Collaborative Knowledge*. MIT Press.
Stegmann, K., Weinberger, A., & Fischer, F. (2007). Facilitating argumentative knowledge construction with computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning*, 2, 421-447.
Toulmin, S. (1958). *The uses of argument.* Cambridge: Cambridge University Press.
Toulmin, S. (1972). *Human Understanding, Volume 1: The Collective Use and Development of Concepts.* Princeton: Princeton University Press.
Windschitl, M. & Hirsch, E.D. (2002). Should Schools Adopt a Constructivist Approach to Education? In Abbeduto, L. (Ed.). *Taking sides: Clashing views on controversial issues in educational psychology*, 2nd edition. Guilford, CT: Dushkin/McGraw-Hill.

Acknowledgments

This material is based upon work partially supported by the National Science Foundation under Grant No. 0822189. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

