Argument as Professional Development: Impacting Teacher Knowledge and Beliefs About Science

Kent J. Crippen

Abstract   Using a case study method, the experiences of a group of high school science teachers participating in a unique professional development method involving an argue-to-learn intervention were examined. The participants ($N = 42$) represented 25 different high schools from a large urban school district in the southwestern United States. Data sources included a multiple-choice science content test and artifacts from a capstone argument project. Findings indicate although it was intended for the curriculum to be a robust and sufficient collection of evidence, participant groups were more likely to use the Web to find unique evidence than to they were to use the provided materials. Content knowledge increased, but an issue with teacher conceptions of primary data was identified, as none of the participants chose to use any of their experimental results in their final arguments. The results of this study reinforce multiple calls for science curricula that engage students (including teachers as students) in the manipulation and questioning of authentic data as a means to better understanding complex socioscientific issues and the nature of science.

Keywords   Argumentation · Professional development · Nature of science

Introduction

When the undergraduate science learning experience of practicing teachers has predominately been lecture and passive involvement, it is not surprising that they struggle with understanding and implementing student-centered inquiry-based strategies. Research has shown that engaging in the process of science as a learner is

K. J. Crippen (✉)
School of Teaching and Learning, University of Florida, 2423 Norman Hall, PO Box 117048,
Gainesville, FL 32611, USA
e-mail: kcrippen@coe.ufl.edu

an important antecedent to using the nature of science as an instructional strategy (Garet et al. 2001; Loucks-Horsley and Matsumoto 1999). Without authentic, practical experience, it would be difficult to envision, let alone enact a strategy unique to an accepted paradigm (Kuhn 1996). One way to address the gap between how science teachers were prepared and their ability to implement the nature of science in their instruction is through professional development. These experiences can become part of a critical intervention for reform by explicitly addressing the content learning of in-service science teachers.

The problem of reforming teacher practice can be compounded when the learning component of professional development is delivered with traditional pedagogies (Luft 2001; Roehrig and Luft 2004). The same outcome can occur even if reformed practices are used, but participants are not held accountable for acquiring a deeper understanding of the subject matter (Singer et al. 2011). Using traditional teacher-centered strategies for delivering science content in professional development contradicts the reform message of inquiry as an appropriate learning strategy (NRC 1996, 2000). Regardless of the delivery strategy, not holding teachers responsible for learning science sends the message that professional development is more about the experience and process than it is about outcomes (e.g., deeper understanding, better practice). These issues constitute a call for reformed professional development in relation to how teachers are engaged in learning science and the degree to which they are held accountable for acquiring a deeper understanding of science.

This study sought to advance theory by evaluating the efficacy of an argue-to-learn professional development experience for a group of high school science teachers related to climate change. The experience was intended to address the need for inquiry-based science learning practices in professional development that produce improved understanding of science content and through personal experience, reformed beliefs about the nature of science. Outcomes of this work included demonstrating the fidelity of argue-to-learn for improving the science content knowledge of in-service teachers. Also, teachers' arguments were suggestive of barriers to achieving beliefs consistent with those engaged in reforming their teaching practice. The relevant literature for this study focuses on the rationale for argument as a teaching and learning strategy for school science.

Review of Related Literature

Argument is a form of reasoning that emphasizes the use of evidence in support of a claim. Toulmin (1958) defines the epistemic elements of an argument as including evidence, claim, warrant, rebuttal, and backing. Driver et al. (2000, p. 293) illustrate the basic structure of Toulmin's argument in sentence form as "because (data)… since (warrant)… on account of (backing)… therefore (conclusion)." In the context of education, the term argument defines the elements used in constructing claims while argumentation describes the process and discourse (Sampson and Clark 2008). Recently, argument and argumentation have become of increasing interest in science education, as a component of the nature of science and as a learning mechanism and pedagogy (Erduran and Jimenez-Aleixandre 2007). While argument

is recognized as the epistemic practice of science, its use as a learning mechanism is a relatively recent endeavor and is termed *Argue-to-Learn*.

Argument is considered a core practice of scientific discourse, one by which scientific data are vetted and eventually accepted as theory (Bricker and Bell 2008). Sandoval (2005) defines argument as an epistemic practice of science that serves a rhetorical function. He proposes that students must be engaged in the activity of developing and defending arguments not just as a skill of science, but as a means of organizing and assigning "epistemic status" to claims (p. 26).

Informal reasoning is the mechanism of argumentation and represents the underlying processes involved in developing a claim. Sadler and Zeidler (2005b) distinguish and differentiate informal reasoning and argumentation according to the following:

> Informal reasoning refers to the cognitive and affective processes involved in the negotiation of complex issues and the formation or adoption of a position. Argumentation refers to the expression of informal reasoning. (p. 73)

A robust understanding of the nature and function of argument implies a mature view of the nature of science (Bell and Linn 2000) and the evaluation of a scientific knowledge claim is accepted by many as a key component of scientific literacy (Wallace et al. 2004). Andriessen (2006) suggests that learning science via argumentation supports the development of a scientific habit of mind and defines arguing to learn as "a form of collaborative discussion in which both parties are working together to resolve an issue, and in which both scientists expect to find agreement by the end of the argument" (p. 443). This type of argument is constructive and reflective, as opposed to more aggressive forms like debate. With argue-to-learn, the goal is a robust understanding built through collaborative and supportive discourse focused on the epistemic components of scientific knowledge: claim, warrant, and evidence.

The utility of argumentation as an instructional and learning model can be justified from philosophical, cognitive, and sociocultural perspectives (Erduran et al. 2004). Learning science with inquiry should involve the accepted mechanisms for establishing scientific knowledge as a distinct and culturally established body of information. As such, teaching science implies the use of argument (Zimmerman 2000). Despite the recent attention that argument has been given, research has not shown it to be a common occurrence in science classrooms (Cavagnetto 2010). Not surprisingly, argue-to-learn is inconsistent with traditional undergraduate science instruction.

Empirical research involving interventions using argue-to-learn has focused primarily on K-12 science students under the guise of reformed teacher practice. Generally, argumentation as an instructional intervention has been shown to improve student conceptual understanding of science concepts, in addition to increasing students' ability to construct arguments (Cross et al. 2008; Venville and Dawson 2010). Recent work has established prior knowledge as a mediator for the effect of an argumentation intervention on the part of students (Cross et al. 2008; Sadler and Zeidler 2005b). Further, Berland and McNeill (2010) make a strong case for the importance of the learning environment for engaging and scaffolding the successful participation of students in the discourse and practice of argumentation.

Fading worked examples, a technique that involves the progressive removal of elements from a richly described solution over time, self-explanation prompts, and other textual supports have shown to be successful for scaffolding student use of argument and scientific explanation (Linn et al. 2005; McNeill et al. 2006; Quintana et al. 2004; Sandoval and Millwood 2005; Schworm and Renkl 2007). It is generally accepted that student argumentation skills develop over time and if scaffolded through professional development, the same is true for teacher use of argumentation as an instructional strategy (Osborne et al. 2004). As an indicator of the current state of research in this area, Berland and McNeill (2010) have proposed a learning progression for scientific argumentation, the first attempt to describe how the conceptual understanding and skills of argumentation might build over time.

While the development of argumentation pedagogy and its impact on students is an active area of research (e.g., Bell and Linn 2000; Osborne et al. 2004), few studies have addressed the potential of argue-to-learn experiences for in-service teachers as a means for developing reformed beliefs and practice. Driver et al. (2000) suggest the importance of teachers in scaffolding student argument. Sadler (2006) recently examined the use of argumentation with pre-service teachers within the context of a methods course, but did not follow-up on the fidelity of the intervention for classroom practice. However, a case study by Avraamidou and Zembal-Saul (2005) documented the role of pre-service teachers' argument experiences on first-year teachers' use of claims and evidence in their science teaching. Simon et al. (2006) developed a professional development program for teachers and have examined how participating teachers progress in their use of argumentation pedagogy.

By examining an intensive argue-to-learn intervention for teachers, this study addresses the lack of research on argumentation experiences as a professional development strategy and answers Sadler and Zeidler's (2005a) call for additional research exploring the reasoning patterns of other target populations by considering the development of in-service teachers.

Study Design

A 2-week summer institute for secondary science teachers was conducted using an argue-to-learn framework, with a focus on developing content knowledge related to the science of global climate change: a current, multidisciplinary, and critical socioscientific issue impacting life on Earth. Sadler and Zeidler (2005a) define socioscientific issues as "open-ended, ill-structured, debatable problems subject to multiple perspectives and solutions." (p. 113). Sadler's (2004) findings concerning student conceptualizations of the nature of science and evidence about global warming as a socioscientific issue included a call for interventions that personalize such issues, address the value of justifying claims, and emphasize counter positions. When applied to classroom teachers as the student audience, the argue-to-learn intervention developed for this study met these criteria.

This study sought to describe how teacher participants used their knowledge of science in constructing a detailed argument related to global climate change. A case study method was used to address the following research questions:

- How did the argue-to-learn intervention impact participant science content knowledge?
- What do the elements of the final arguments indicate about participant reasoning in relation to the efficacy of argue-to-learn for science?

Improving science content knowledge was the primary goal for the professional development and is the rationale that supports research question one. Pragmatically, it was difficult to promote argue-to-learn without first assessing its utility for producing measurable learning. Further, it was assumed that demonstrating a measurable learning gain for the participants who were also science teachers might provide added motivation for personally adopting this instructional strategy.

The Argue-to-Learn Intervention

The argue-to-learn intervention took the form of an 8-day, 6-h per day formal educational experience as a summer science institute. Teacher participants earned 3 credits of graduate-level science credit for successfully completing the institute. The curriculum was geared toward the construction of assigned arguments and instruction took the form of alternating whole group and break out room sessions. The whole group sessions involved traditional lectures that were designed to provide information about using the Toulmin model for constructing scientific arguments and to hear expert opinion regarding the science of climate change. The breakout room sessions were led by peer facilitators and were intended to engage small groups in activities to explore the science of climate change as evidence, to scaffold the development of arguments with feedback, and to offer low-stakes, informal opportunities to engage in argument. Within the breakout room sessions, participants spent time discussing how evidence could be used to support claims from different perspectives while making explicit the reasoning involved in doing so. The argument scaffolds included evidence-claim-reason statements as narrative text, tree diagrams, and the graphical elements of hypothesis, data, and relationship used by the mapping software. In order to keep the argument sessions constructive and collegial, participants were held responsible for acceptable behaviors for argumentation, including making thinking explicit, being critical of ideas and not people, and citing evidence known to the group (Chinn 2006).

Using methods reported in the literature, the learning materials of the summer institute were adapted from traditional materials to enact argue-to-learn (Kuhn and Reiser 2006; Osborne et al. 2004). Three existing, commercially available probeware activities as well as a field trip to a local nature preserve were adapted using the Science Writing Heuristic (SWH) (Hand and Keys 1999). This adaptation involved parsing the directions from the cookbook form of the activity into a description and directions for the procedure section of the SWH form. Specific to each activity, questions and writing prompts were added to cue beginning ideas and to make a prediction about the pending graph. Data were collected in tabular and graphical forms and participants were prompted to use this as evidence for constructing an appropriate claim. These evidence-claim statements were compared and critiqued among groups, then compared to the assigned reading materials. The

final component included a reflection prompt to describe how their ideas may have changed as a result of the inquiry and to generate new questions.

In addition to the experiments, the curriculum included a range of activities that served as potential evidence for arguments (Table 1). The intent was to provide a robust, manageable set of materials that expressed both sides of the climate change debate. This material was tailored for use in arguments, so that participants would not have to spend their time identifying additional evidence. This intent was made explicit on the first day of the institute, but participants were not discouraged or denied access to the Web. In fact, a Web-based course management system was used to manage all of the institute materials electronically. Wireless access and a small collection of computers were provided, and participants were encouraged to bring their personal computers. All assignments and materials were made available initially and continuously throughout the institute. Participants were expected and encouraged to have been continuously revising their arguments as the institute progressed.

As a capstone project, participants were responsible for presenting a group argument and creating an electronic argument map. Groups of participants were randomly assigned to one of six arguments (Appendix). These arguments took the form of either a pro or con statement related to three themes: the model for climate change, a need for additional data, and a need for exploring other important questions. Previous research has established these themes as useful for engaging learners in the science content of a controversial issue (Passmore 2007). The learning goal of the activity was to prepare and present the best argument possible in support of your position, while also addressing the counter argument that used the curriculum as evidence. Electronic argument maps were completed individually as homework and each group made a time-limited presentation of their argument and responded to questions. Participants were given basic instructions on using the software application Belvedere (version 4.1) to construct their argumentation maps (Toth et al. 2002). The form of the argument was not constrained by requirements of the project or our use of Toulmin's model, and participants were given complete freedom to create what they felt was appropriate. Participants who were not part of the presentation team completed an evaluation rubric for each group (Toth et al. 2002) and everyone was provided a copy of the rubric on the first day.

Participants were first educated on the form and function of argument and then required to use the structure for processing the content of the summer institute. On the first day of the summer institute, participants were introduced to Toulmin's model (1958) and given background information on how the use of argument was built into the activities (Voss and Means 1991). Three example arguments were presented and discussed. Finally, the capstone project was explained and participants were randomly assigned to their argument and groups.

Participants

Research participants were secondary science teachers ($N = 42$) involved in a 2-week summer science institute that was a component of a long-term professional development project (Crippen et al. 2010). The participants included 27 females,

Table 1 A chronology of the structured curriculum elements, delivery format, ideas presented, and intended relevance for argue-to-learn

Element	Title	Format	Materials used	Ideas presented	Relevance for argue-to-learn
Reading	The Nature of Argumentation	Ind	pp. 295–301 of Driver et al. (2000)	Overview of argue-to-learn	Background knowledge and reflection
Reading	Climate Change 101	Ind	Pew (2007)	Basic ideas of climate change science	Background knowledge, evidence, warrants, and argument examples
Lecture	The Essence of Argumentation	Lg Grp	Slideshow	Nature of argumentation in science, Toulmin's model, capstone project	Background knowledge, details of the intervention, and expected outcomes
Web resources	Recent Evidence for Climate Change	Ind	25-Articles from Science News	Results of scientific research	Empirical scientific sources of evidence
Panel	Perspectives on Climate Change	Lg Grp	Oral presentation	Multiple perspectives and nature of evidence	Evidence warrants and argument examples
Lecture	Lab Work with the Science Writing Heuristic (SWH)	Lg Grp	Slideshow	Key elements of SWH, relation to traditional lab reports	Working within an argument scaffold for evidence-claim statements
Lab activity	The Greenhouse Effect	Sm Grp	Temperature sensor SWH activity	Environments are created with one spiked with CO_2 gas. The temperature was monitored over time	Developing/critiquing a lab evidence-claim statement. Evidence and warrants
Lecture	Teaching with Analogies	Lg Grp	Slideshow	Nature of analogical reasoning, implications for learning & teaching	Analogy as a form of warrant
Speaker	A Geologist's Perspective on Climate Change	Lg Grp	Slideshow	Geologic/astronomical processes underlying human induced change	Evidence, warrants, and argument examples
Discussion	Collaborative Argument Review and Critique	Sm Grp	Assigned arguments	Presentation of arguments, review of evidence, and consideration of rebuttal	Argument practice. Formative peer feedback
Lab activity	Sea Flooring Spreading	Sm Grp	Magnetic field sensor SWH activity	The movement of plates is monitored for a model of the ocean floor	Developing/critiquing a lab evidence-claim statement. Evidence and warrants

Table 1 continued

Element	Title	Format	Materials used	Ideas presented	Relevance for argue-to-learn
Field trip	Nature Preserve	Lg Grp	Guided tour	The ecology of the area, the impact of warming, and sustainable living	Evidence, warrants, and argument examples
Lab activity	Climate at the Preserve	Sm Grp	Temperature, light, UVA/UVB sensors SWH activity	Develop and complete an open inquiry on climate	Developing/critiquing a lab evidence-claim statement. Evidence and warrants
Video	An Inconvenient Truth	Ind	DVD rental	Pro argument for human impact on climate	Evidence, warrants, and argument examples
Lecture	Analogies for Climate Change	Lg Grp	Slideshow	Analogies explaining climate change	Examples of using an analogy to warrant a claim
Speaker	A Meteorologist's Perspective on Climate Change	Lg Grp	Slideshow	Local temp change, human influence, and data issues	Evidence, warrants, and argument examples
Lab activity	Respiration	Sm Grp	CO_2 gas and temperature sensor SWH activity	The respiration of crickets was measured with temperature	Developing/critiquing a lab evidence-claim statement. Evidence and warrants
Web activity	Climate Change, Who's to Blame?	Sm Grp	Web-based inquiry science environment (WISE)	The human impact on climate issue in scaffolded argument form	Developing/critiquing an evidence-claim statement. Evidence, warrants, and practice arguing
Video	The Great Global Warming Swindle	Lg Grp	Digital copy	Con argument for human impact on climate	Evidence, warrants, and argument examples
Speaker	A Biologist's Perspective on Climate Change	Lg Grp	Slideshow	The reciprocal relationship between climate and life	Evidence, warrants, and argument examples
Lecture	Using Analogies in the Classroom	Lg Grp	Slideshow	Best practice	Evidence, warrants, and argument examples

Ind, individual; Sm Grp, small group; Lg Grp, large group

15 males, and represented 25 different high schools from a large urban school district in the southwestern United States.

Method

The methodology for this study took the form of a single case study design (Creswell 1997). The case was bounded by the intent to understand the experience of a group of high school science teachers participating in a unique professional development that provided them with the experience of learning science via argue-to-learn. In this case, argue-to-learn involved using evidence to warrant a scientific claim concerning human-induced climate change and defending a rebuttal. The experience was designed to explicitly emphasize argument as a means to understanding scientific concepts and as an appropriate pedagogy for science education. The limited research involving argue-to-learn for in-service teachers dictated the context-dependent method and inductive data analysis.

Data Sources

Data sources included a multiple-choice science content test, electronic argumentation maps, digital video of the final group presentation, artifacts used in the final presentation, and evaluation matrix scores for the final group presentation. The final piece of data was a semi-structured focus group interview with a smaller, convenience sample of participants. The group interview addressed the validity of the themes developed in the analysis of the argument artifacts and explored the underlying causal factors.

Results

Science content knowledge was evaluated using a survey instrument consisting primarily of items from existing concept inventories; a recent method of assessing domain knowledge for undergraduate students (Evans and Hestenes 2001). A 25-item multiple-choice instrument was developed, using questions from two different concept inventories and a commercially available test question software program: The Greenhouse Effect Concept Inventory (Keller 2006), the Geosciences Concept Inventory (Anderson and Libarkin 2006; Libarkin and Anderson 2005), and ExamGen (2007). A paired-samples t test was conducted on the participants' scores and a statistically significant increase was found from pretest ($M = 15.16$, SD $= 5.10$) to posttest ($M = 17.30$, SD $= 3.85$), $t_{42} = -5.442$, $p < 0.001$, $\eta^2 = 0.41$. The normalized gain for all participants was 0.18, which is considered a low gain.

Using a constant comparative method (Lincoln and Guba 1985), a content analysis was conducted on the argumentation project materials focused on the type of resources and manner in which they were used as elements of Toulmin's model. To address research question 2, the analysis emphasized the content of the arguments rather than the process. Participants were asked to use the curriculum as

evidence in their arguments, so initial codes were developed from the major activities (e.g., speakers as evidence, warrant by fact) and additional codes were added during analysis (e.g., K-12 curriculum materials and analogy as evidence). A warrant was operationally defined as the stated reason why a form of evidence supported a claim and participants were given minimal guidance on how to create them. Thus, the categories for types of warrants were emergent from the data sources and refined through constant comparison. After the data were classified and coded, the raw count of codes were normalized as percentages and compared descriptively.

A total of 391 sources of evidence were cited by participants and coded into 13 categories (Table 2). Coding involved studying the reference for each evidence citation and comparing it to the attributes of the existing categories to determine goodness of fit. Evidence included inscriptions in the form of charts and graphs that were created by participants and used during their final presentations. Roth and McGinn (1998) define inscriptions as "signs that are materially embodied in a medium" (p. 37). These inscriptions may have been re-creations or copies of other primary sources, but no reference was provided for the source. As such, these were coded into the category of "no reference." Without a citation for a primary source, it was not possible to assess the accuracy of the reproduction of the artifact or the science they were intended to communicate. The "outside reference" category included inscriptions that were not provided to participants, but were cited as being discovered through use of the Web.

Participant use of evidence produced interesting and unintended results. Contrary to the designed intent of 100 %, only 25 % of the evidence used came from provided resources. In comparison, the number of outside resources was greater than the combined use of all provided resources. Based largely on the continuing positive response expressed in formal evaluation measures, university science faculty as guest speakers is a primary component of our professional development model (Crippen et al. 2010). However, in all arguments, participants made little reference to the opinions of these speakers as evidence. Interestingly, not one group or individual participant used the results of the laboratory experiments they conducted as evidence.

A total of 847 warrants were coded into five categories (Table 3). The warrant category of *Fact* often included the declaration of multiple fact statements followed immediately by a claim statement. Rarely did the participants make explicit their reasoning for how a particular piece of evidence supported a claim; instead, they left this task to be discerned by the audience or reader. Coding involved interpreting the reasoning of participants by carefully examining first, the implied relationship among collections of evidence and then between the collection or individual piece and the claim statement (Fig. 1).

As documented in the number of pieces of evidence cited and the number of warrants used in the arguments, the nature of the assigned argument influenced the participant's response. In all cases, more sources of evidence were provided for the pro side of all arguments. With the exception of the argument related to important questions (set C, 5&6), the pro side of every argument had a greater number of warrants. Generally, the pro side of arguments (1, 3, 5) were composed of warrants

Table 2 A comparison of sources of evidence by question type used in all of the final argument projects

	Overall	1. Human cause	2. Human not cause	3. More research	4. Less research	5. New questions	6. Same questions
No reference	154 (39%)	32 (39%)	34 (58%)	24 (31%)	16 (29%)	16 (26%)	32 (59%)
Outside reference (not provided)	104 (27%)	21 (25%)	3 (5%)	36 (46%)	22 (39%)	17 (28%)	5 (9%)
Provided references (document or URL)	40 (10%)	18 (22%)	1 (2%)	1 (1%)	0 (0%)	13 (21%)	7 (13%)
Pre-reading (Pew 2007)	19 (5%)	0 (0%)	13 (22%)	2 (3%)	4 (7%)	0 (0%)	0 (0%)
Previously cited data	17 (4%)	1 (1%)	6 (10%)	2 (3%)	4 (7%)	0 (0%)	4 (7%)
Analogy—new, created by participant	16 (4%)	2 (2%)	2 (3%)	0 (0%)	9 (16%)	2 (3%)	1 (2%)
Video—great global warming swindle	12 (3%)	0 (0%)	0 (0%)	11 (14%)	0 (0%)	1 (2%)	0 (0%)
Speaker—national weather service	7 (2%)	4 (5%)	0 (0%)	0 (0%)	1 (2%)	1 (2%)	1 (2%)
Speaker—geologist	6 (2%)	0 (0%)	0 (0%)	1 (1%)	0 (0%)	3 (5%)	2 (4%)
Speaker—biologist	5 (1%)	4 (5%)	0 (0%)	0 (0%)	0 (0%)	1 (2%)	0 (0%)
K-12 curriculum materials	4 (1%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	4 (7%)	0 (0%)
Video—an inconvenient truth	3 (1%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (2%)	2 (4%)
Analogy—provided, created by facilitators	4 (1%)	1 (1%)	0 (0%)	1 (1%)	0 (0%)	2 (3%)	0 (0%)
	391	83	59	78	56	61	54

Table 3 A comparison of warrants by question type used in all of the final argument projects

|	Overall	1. Human cause	2. Human not cause	3. More research	4. Less research	5. New questions	6. Same questions
Fact (fact statement(s), then new claim)	413 (49%)	45 (49%)	14 (32%)	53 (55%)	33 (43%)	42 (53%)	42 (51%)
Assumed logical step (statement of logic; often multiple claims, then new claim)	292 (34%)	31 (34%)	28 (64%)	26 (27%)	35 (45%)	20 (25%)	21 (25%)
No evidence (new claim only)	72 (9%)	8 (9%)	1 (2%)	9 (9%)	7 (9%)	5 (6%)	14 (17%)
Amount of evidence (description of multiple sources of evidence; focus on quantity, then new claim)	42 (5%)	0 (0%)	1 (2%)	7 (7%)	2 (3%)	10 (13%)	2 (2%)
Expertize (as stated by {Expert Name}, then new claim)	28 (3%)	8 (9%)	0 (0%)	1 (1%)	0 (0%)	3 (4%)	4 (5%)
Total	847	92	44	96	77	80	83
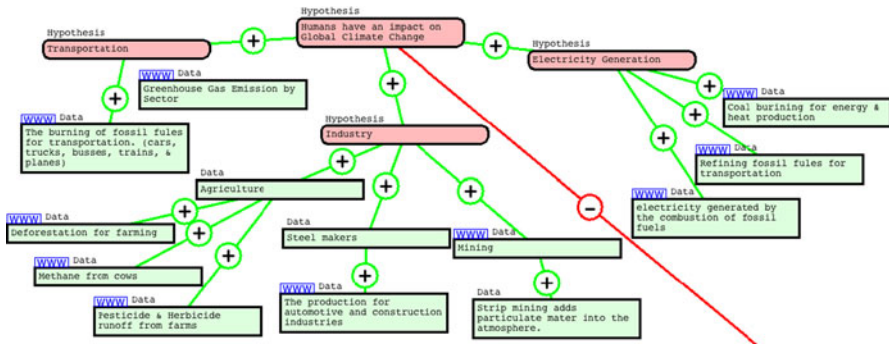
Fig. 1 Part of a participant argument map that was built to address argument-1, humans are causing global climate change. This example illustrates the use of outside Web resources as references (*labeled* as WWW) and warrants of assumed logical step and amount of evidence

by fact for outside references as evidence, while the con side (2, 4, 6) had a greater occurrence of warrant by logical step for non-referenced evidence. The arguments related to humans as the underlying cause of climate change (Set A, #1&2) generated the greatest citations of evidence while the arguments related to additional data (Set B, #3&4) generated the greatest number of warrants.

From the very beginning until the end, participants expressed uneasiness with the notion of warrant. While teachers seemed comfortable defining the concept and identifying examples, they struggled with constructing their own models. The limited number of types as well as complexity of the warrants may be due to the difficulty participants experienced in understanding the concept. However, when prompted to explain reasoning during informal practice sessions, participants often replied that they assumed that the warrant was obvious, indicating that the reasoning was already known by all and thus did not need to be formally stated. Though not explored formally here, these data may represent a misconception or be indicative of other issues with content understanding or efficacy.

Finally, the group presentations were judged by the remaining participants as an audience using the rubric provided by Toth et al. (2002). This 11-item survey instrument included such items as "Supporting data is listed for each argument." and "The team's interpretation of data is accurate." Participants responded on a 3-point scale from *1-not at all* to *3-always*. These raw scores were averaged and the highest cumulative average score was used to determine the best argument. The scores were not reported until 1 day following the conclusion of the summer institute and had no bearing on participant grades.

Though the audience rated all of the arguments relatively high, the types of evidence and warrants used in the more highly rated arguments were different from what was used in the less highly rated arguments. The argument for no human impact on climate change was the highest rated argument ($N = 38$, $M = 30.8$, $SD = 2.01$) and used a significant amount of non-referenced evidence with assumed logical warrants (Tables 2, 3). The arguments related to the theme of the need for additional data (i.e., propose an additional study) received the lowest final rubric evaluation scores ($N = 38$, $M = 27.5$, $SD = 3.87$). These two groups used the

highest percentage of outside resources as evidence and used fewest of the provided references.

Analysis of these data produced the following themes that were member checked on two separate occasions: (a) instead of using the available resources to theorize, participants spent their time searching the Web to identify compelling inscriptions, (b) participants were uncomfortable and unsure how to use their personal, experimental results in their arguments, and (c) the "best" arguments were those that used many facts (often presented as an inscription) to support a claim via an assumed logical warrant (i.e., it was not necessary to explain explicitly how each piece of evidence supported a claim; "it's obvious, we're all science teachers").

Nine months after the summer institute, a focus group of participants ($N = 11$) was convened to explore the themes from the content analysis. The focus group was presented with the results, and the themes were used as a semi-structured interview guide. Data were recorded as field notes, compared against the content analysis, and member checked with a single participant.

Participants accounted for the lack of laboratory results as evidence by expressing an apprehension or lack of "trust" in their ability to generate data that could be used as primary evidence for a scientific argument. This uneasiness was personal and seemed to be rooted in issues of accuracy and reliability. Even though other participants reproduced their results independently, they expressed a strong desire to repeat the experiments with a greater attention to detail and repeated trials. In addition, their unfamiliarity with probeware may have contributed to the feelings of discomfort. This finding is consistent with previous research where teachers had issues distinguishing data from scientific knowledge (Taylor and Dana 2003).

The unanticipated use of outside resources as evidence (especially those identified through Web searching) was attributed to a need to find evidence that more directly supported their positions. Their descriptions portray a process that valued time spent looking for evidence to fit a position, as opposed to time spent evaluating existing data and developing warrants. From the beginning, participants were uncomfortable and unclear about the notion of warrant. They assumed that representations of data (e.g., charts and graphs) warranted claims without need for explanation. Thus, they viewed their job in preparing their arguments as one of finding graphs with surface features that supported their positions. Further, some described the need for "the element of surprise" by having yet unseen references for their presentations.

Discussion

For the group of in-service high school science teachers in this case study, the argue-to-learn intervention was successful for improving their science content knowledge. Though this gain was statistically significant, the magnitude of change was not large. From the data collected, it is unclear whether the gain is limited by the intervention, the survey instrument, or a ceiling effect due to the general and interdisciplinary nature of the topic. However, considering the need for reformed methods of professional development that impact content knowledge as well as

pedagogical knowledge, skills, and dispositions, this result is important and encouraging.

The resources and experiences from the institute were not translated into the final arguments as they were designed. Though it was intended for the curriculum to be a robust and sufficient evidence collection for supporting any of the arguments, participant groups were more likely to use the Web to find unique evidence, in the form of undocumented inscriptions. While this finding was unanticipated, it does support the assertion that one of the primary roles of inscriptions is to serve as social objects for supporting argumentation (Roth and McGinn 1998). Further, this finding is consistent with Sandoval and Millwood (2005) who found that students' references to specific inscriptions often failed to articulate how specific data related to particular claims. These unwarranted assumptions could be interpreted as similar to typical issues that have been reported for student arguments (Toulmin et al. 1984).

Some of the findings are consistent with existing theoretical assumptions as well as other reported findings. For example, the familiarity of the arguments related to humans as the underlying cause of climate change as well as the arguments related to additional data may have accounted for the greater complexity of the participants' response. Bricker and Bell (2008) suggest that familiarity with a subject affects argumentation practice and produces more complex linguistic and cognitive processes. According to Andriessen (2006), the trend for participants to create more complex arguments for the pro side of any issue is a common finding. The lack of use of content experts as evidence is consistent with previous research suggesting that people tend to accept expert scientific opinion, but do not use them in future argumentation activities (Schwarz and Glassner 2003). If the heavy use of outside references is viewed as a goal directed activity focused on constructing the best corpus of evidence, then that activity is consistent with the findings of Choi et al. (2010), who found that argument quality for older students was best predicted by the evidence they constructed. Finally, consistent with Andriessen (2006), the use of argument maps seemed to influence the discourse of participant presentations by focusing the audience members on evidence not presented. "The argument maps became collectively shared scaffolds that allowed students to compare interpretations of evidence." (Andriessen 2006, p. 454).

Perhaps the most intriguing result of this study is that not one of the participants chose to use any of their personal experimental results in their arguments. The focus group was surprised by this finding and ultimately explained it as a lack of trust. However, the technology and nature of the experiments may have confounded the problem. The focus group described the experiments as too simplistic, lacking in rigor, and not the kinds of things that scientists would use to generate complex theories. While all of these statements were plausible, a sufficient response was not provided for the rebuttal question, what types of experiments would need to be done? Remarks focused on the need for advanced instruments, careful measurements, and reproducibility. Their comments and body language indicated that unless they were reproducing a well-known experiment, they were quite uncomfortable using any of their own experimental results for theorizing, especially when presenting to peers. The possibility exists that these teachers did not view the activities as authentic science and may have simply perceived them as a set of

activities that demonstrated scientific concepts. Thus, because the activities were not perceived as authentic, the participants could not rationalize their use because they did not qualify as data that could be used to support scientific arguments.

The problem of using primary data for theorizing may be connected to past student experiences in science. If they had not previously learned this skill, then they have no prior knowledge or experience to draw from in this case. Consider that Sadler, et al. (2004) found that approximately one-half of the 84 high school students in their sample were not able to accurately identify and describe data. If these same students were to continue their schooling and eventually become science teachers and were never asked to theorize from primary data, then their knowledge and skills in this regard may mimic those of the study participants. Finally, the lack of using empirical data may support Sandoval's (2005) notion about the influence of a practical epistemology in the practice of inquiry and how it might create differential results from the application of a formal epistemology.

The warrants used in the arguments were more implicit than explicit. Participants were more likely to state a string of facts and then make a claim, rather than to explain how each fact might directly support the claim. In designing the intervention, it was assumed that participants would be motivated intrinsically by interest and the need to know and extrinsically by their inclusion within a professional peer group to make better sense of evidence in light of specific scientific claims and to seek a deeper explanation for results. However, there seemed to be a general sense that they were among knowledgeable others, so there was no need to explain their reasoning. This finding could be attributed as a wrong assumption due to the teachers' lack of comfort or efficacy for professional development, but it also could be justified as a strategy for hiding their lack of knowledge. If the latter, then it could explain the shallow content knowledge gains and be perceived as an element of the intervention needing further research and development.

In addition to teacher issues, the materials and design of the intervention may have contributed to the form of the participant arguments. For example, the content of the materials was unintentionally weighted toward the debate about human-induced climate change and this may have been responsible for the extra resources used in these arguments. While the Science News articles encompassed an entire year of the magazine, their topic focus may not have been rich enough to support the needs of all arguments. This was an initial attempt at argue-to-learn for professional development and the instructional methods used may have been too implicit, vague, or poorly delivered. Finally, our limited experience and developing understanding about delivering argue-to-learn professional development could have unknowingly constrained the participants and their arguments.

For this group of participants, the argue-to-learn intervention was successful for building content knowledge. Participants found argumentation to be engaging and fun. However, for some, learning to argue may have become more important than learning science. Clearly, the potential existed for a shift in focus from demonstrating science content knowledge as the priority to being perceived as a good game player. As such, using argue-to-learn requires unique scaffolding and conscience effort by the instructors in order to keep the focus on the science content to be learned.

Conclusion

This study addresses the lack of research on argumentation as a professional development method for improving teacher content knowledge and as a proxy for reformed beliefs about science. The results establish argumentation as a successful method for building teacher content knowledge and engaging them in an authentic science experience. Consistent with the findings of Taylor and Dana (2003), issues were established with teacher conceptions of primary data. Of particular interest was the lack of "trust" they expressed for using their laboratory results for scientific theorizing. This finding calls into question their beliefs about the role of experiments and laboratory experiences for learning and indicates that these beliefs may be an important barrier to reformed practice.

Results of this study have informed the ongoing local use of argue-to-learn for professional development through modifications to the intervention. Current institutes include a more explicit emphasis on the argue-to-learn framework, including a discussion about the findings from previous studies, such as this one. Modifications include a greater number of experiments that require data collection and analysis, a requirement that all participants use the results of these experiments in creating their arguments, and advanced strategies such as Gowin and Alvarez's (2005) Vee Map to scaffold the use of argument. Our current research efforts focus on evaluating the impact of different forms of learning scaffolds as well as exploring the relationships among the intervention, teacher beliefs, and classroom practice.

The results reinforce multiple calls for science curricula that engage students (including teachers as students) in the manipulation and questioning of real data as a means to better understanding complex socioscientific issues as well as the nature of science (Sadler 2004; Tytler et al. 2001). Finally, this study also supports Sandoval's (2005) call for research examining practical epistemological ideas in the context of inquiry practice. In particular, the artifacts generated during the inquiry as well as the supporting discourse.

Appendix: Final Project: Assigned Group Arguments

Set A: Model

1. Propose a model that best supports the theory that humans are impacting global climate change.
2. Propose a model that best refutes the theory that humans are impacting global climate change.

   Set B: Additional data

3. The National Science Foundation has created a highly competitive grant program for research related to global climate change. The amount of money

available is large, but they are only supporting a small number of high impact projects. Develop a research proposal that includes a single study to address the debate related to human impact on global climate change.

4. Congress is debating whether to divert money earmarked for research at the National Science Foundation related to global climate change. Take the position that enough research has been done and that these resources are better allocated for programs intended to solve the global climate change issue.

   Set C: Important questions

5. Many people question the time and resources spent on the issue of "Are humans impacting the environment?" Propose an alternative question or set of questions for science to bring to the forefront of the national policy debate.

6. Identifying the impact of humans on global climate change is paramount to decisions related to current national policy. Take the position that additional research is needed.

References

Anderson, S. W., & Libarkin, J. C. (2006). *The Geoscience Concept Inventory*. Retrieved April, 27, 2007, from http://newton.bhsu.edu/eps/gci.html.

Andriessen, J. (2006). Arguing to learn. In K. Sawyer (Ed.), *Handbook of the learning sciences* (pp. 443–459). Cambridge: Cambridge University press.

Avraamidou, L., & Zembal-Saul, C. (2005). Giving priority to evidence in science teaching: A first-year elementary teacher's specialized practices and knowledge. *Journal of Research in Science Teaching, 42*(9), 965–998.

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), 797–817.

Berland, L. K., & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education, 94*(5), 765–793.

Bricker, L. A., & Bell, P. (2008). Argumentation from science studies and the learning sciences and their implications for the practices of science education. *Science Education, 92*(3), 473–498.

Cavagnetto, A. R. (2010). Argument to foster scientific literacy. *Review of Educational Research, 80*(3), 336–371.

Chinn, C. A. (2006). Learning to argue. In A. M. O'Donnell, C. E. Hmelo-Silver, & G. Erkens (Eds.), *Collaborative learning, reasoning, and technology*. Mahawah, NJ: Erlbaum.

Choi, A., Notebaert, A., Diaz, J., & Hand, B. (2010). Examining arguments generated by year 5, 7, and 10 students in science classrooms. *Research in Science Education, 40*(2), 149–169.

Creswell, J. W. (1997). *Qualitative inquiry and research design: Choosing among five traditions* (3rd ed.). Thousand Oaks: Sage.

Crippen, K. J., Biesenger, K. D., & Ebert, E. E. (2010). Using professional development to achieve classroom reform and science proficiency: An urban success story from Southern Nevada USA. *Professional Development in Education, 36*(4), 637–661.

Cross, D., Taasoobshirazi, G., Hendricks, S., & Hickey, D. T. (2008). Argumentation: A strategy for improving achievement and revealing scientific identities. *International Journal of Science Education, 30*(6), 837–861.

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

Erduran, S., & Jimenez-Aleixandre, M.-P. (Eds.). (2007). *Argumentation in Science Education*. Berlin: Springer.

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

Evans, D. L., & Hestenes, D. (2001, 2001). *The concept of the concept inventory assessment instrument.* Paper presented at the Frontiers in Education Conference, 2001. 31st Annual, Reno, NV.

ExamGen (2007). *ExamGen: Test item databases for today's teachers.* Retrieved June, 14, 2007, from http://www.examgen.com/.

Garet, M. S., Porter, A. C., Dismone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*(4), 915–945.

Gowin, B. D., & Alvarez, M. C. (2005). *The art of educating with V diagrams.* Cambridge: Cambridge University Press.

Hand, B., & Keys, C. W. (1999). Inquiry investigation: A new approach to laboratory reports. *The Science Teacher*(April), 27–29.

Keller, J. (2006). *Eliciting and addressing student misconceptions regarding the atmospheric greenhouse effect and radiative equilibrium.* Arizona: University of Arizona.

Kuhn, T. S. (1996). *The structure of the scientific revolution* (3rd ed.). Chicago, IL: University of Chicago Press.

Kuhn, L., & Reiser, B. (2006). *Structuring activities to foster argumentative discourse.* Paper presented at the American Educational Research Association, San Francisco, CA.

Libarkin, J. C., & Anderson, S. W. (2005). Assessment of learning in entry-level geoscience courses: Results from the Geoscience Concept Inventory. *Journal of Geoscience Education, 53*(4), 394–401.

Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry.* Thousand Oaks, CA: Sage Publications.

Linn, M. C., Bell, P., & Davis, E. A. (2005). *Internet environments for science education.* London: Lawrence Erlbaum.

Loucks-Horsley, S., & Matsumoto, C. (1999). Research on professional development for teachers of mathematics and science: The state of the scene. *School Science and Mathematics, 99*(5), 258–271.

Luft, J. A. (2001). Changing inquiry practices and beliefs: The impact of an inquiry-based professional development programme on beginning and experienced secondary science teachers. *International Journal of Science Education, 23*(5), 517–534.

McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *Journal of the Learning Sciences, 15*(2), 153–191.

NRC. (1996). *National science education standards.* Washington, DC: National Academy Press.

NRC. (2000). *Inquiry and the national science education standards: A guide for teaching and learning.* Washington, DC: National Academy Press.

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

Passmore, C. M. (2007). *Argumentation in modeling classrooms.* Paper presented at the National Association for Research in Science Teaching, New Orleans, LA.

Pew (2007). *Climate Change 101: Understanding and Responding to Global Climate Change.* Retrieved 12/5, 2010, from http://www.pewclimate.org/global-warming-basics/climate_change_101.

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

Roehrig, G. H., & Luft, J. A. (2004). Constraints experienced by beginning secondary science teachers in implementing scientific inquiry lessons. *International Journal of Science Education, 26*(1), 3–24.

Roth, W.-M., & McGinn, M. K. (1998). Inscriptions: Toward a theory of representing as social practice. *Review of Educational Research, 68*(1), 35–59.

Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching, 41*(5), 513–536.

Sadler, T. D. (2006). Promoting discourse and argumentation in science teacher education. *Journal of Science Teacher Education, 17*(4), 323–346.

Sadler, T. D., Chambers, W. F., & Zeidler, D. L. (2004). Student conceptualizations of the nature of science in response to a socioscientific issue. *International Journal of Science Education, 26*(4), 387–409.

Sadler, T. D., & Zeidler, D. L. (2005a). Patterns of informal reasoning in the context of socioscientific decision making. *Journal of Research in Science Teaching, 42*(1), 112–138.

864

...

Sadler, T. D., & Zeidler, D. L. (2005b). The significance of content knowledge for informal reasoning regarding socioscientific issues: Applying genetics knowledge to genetic engineering issues. *Science Education, 89*(1), 71–93.

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

Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education, 89*(4), 634–656.

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

Schwarz, B., & Glassner, A. (2003). The blind and the paralytic: Supporting argumentation in everyday and scientific issues. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 227–260). Dordrecht: Kluwer.

Schworm, S., & Renkl, A. (2007). Learning argumentation skills through the use of prompts for self-explaining examples. *Journal of Educational Psychology, 99*(2), 285–296.

Simon, S., Erduran, S., & Osborne, J. (2006). Learning to teach argumentation: Research and development in the science classroom. *International Journal of Science Education, 28*(2), 235–260.

Singer, J., Lotter, C., Feller, R., & Gates, H. (2011). Exploring a model of situated professional development: Impact on classroom practice. *Journal of Science Teacher Education, 22*(3), 203–227.

Taylor, J. A., & Dana, T. M. (2003). Secondary school physics teachers' conceptions of scientific evidence: An exploratory case study. *Journal of Research in Science Teaching, 40*(8), 721–736.

Toth, E. E., Suthers, D. D., & Lesgold, A. M. (2002). Mapping to know: The effects of representational guidance and reflective assessment on scientific inquiry. *Science Education, 86*(2), 264–286.

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

Toulmin, S., Rieke, R., & Janik, A. (1984). *An introduction to reasoning* (2nd ed.). New York, NY: Macmillan.

Tytler, R., Duggan, S., & Gott, R. (2001). Dimensions of evidence, the public understanding of science and science education. *International Journal of Science Education, 23*(8), 815–832.

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

Voss, J. F., & Means, M. L. (1991). Learning to reason via instruction in argumentation. *Learning and Instruction, 1*(4), 337–350.

Wallace, C. S., Hand, B., & Yang, E.-M. (2004). The science writing heuristic: Using writing as a tool for learning in the laboratory. In E. W. Saul (Ed.), *Crossing borders in literacy and science instruction: Perspectives on theory and practice* (pp. 355–368). Newark, DE: International Reading Association.

Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental Review, 20*(1), 99–149.