Electronic Science Notebooks and Argumentation: Analysis of Student Writing

Objectives or Purpose

Inquiry-based learning is a part of modern science education at all levels. The new science education standards framework has both emphasized the central role of inquiry and further defined the core competencies, such as making a claim and supporting it with evidence (NRC, 2011). This emphasis on argumentation arises not only out of its importance in science practice but also because students tend to struggle with mastering it (McNeill, 2011), especially at the elementary level. However, when elementary students are provided with the appropriate scaffolds and supports, researchers report that they are able to engage in and develop the core elements of argumentation (Varelas et al., 2008).

To this end, the research team developed an electronic science notebook (Leonardo Project, 2011), the CyberPad, to provide a rich, structured inquiry environment where students would be able to write and draw about their experiences as they were guided through classroom activities by both their teacher and the CyberPad. The CyberPad units were designed to be used in conjunction with the FOSS™ kit-based curricula (FOSS Project, 2008) for this reported work, specifically the Energy and Electromagnetism unit. Prompts were written that encouraged students to both engage in the provided science conceptual knowledge and observations of science phenomena, and develop arguments around both predictions and conclusions of the inquiry lab activities.

The purpose of this reported research is to explore the extent to which fourth grade students were able to engage in meaningful written argumentation when prompted to do so. Of specific interest is the degree to which students chose to provide only claims or evidence (i.e., justifications) or both and the nature of the evidence provided (McNeill, 2011; Sampson & Clark, 2008). Prompts were provided prior to making observations (i.e., collect data) where their justifications would more likely have been hypothetical, and also after observations for the investigation where students would have the opportunity to engage in empirically based justifications.

Perspective or Theoretical Framework

This research works under the assumption that having students construct scientific explanations and engage in discursive argumentation is an essential part of classroom-based science learning and instruction (Berland & Reiser, 2008; Driver, Newton, & Osborne, 2000). For that reason, it is important to be able to identify and assess student argumentation for the purposes of both understanding their ability to think scientifically and designing effective instruction (cf., McNeill & Krajcik, 2008). The basic form of Toulmin's (1958) argumentation structure has been interpreted and adopted by numerous science education researchers in order to assess the substance and quality of discursive arguments that students construct during inquiry-oriented investigations (Sampson & Clark, 2008). Of particular interest is the ability to identify claims that students make and the evidence (i.e., justification) that they may also provide to support the claim. Using a framework similar to one employed by Berland and Reiser (2008), evidence can be either empirical, data collected as part of the investigation or hypothetical, a conjecture (e.g., a prediction) based on another source of knowledge, such as material presented in class, read about on the CyberPad, or learned outside of the classroom. Empirical evidence of specific events that students cite can also be logically linked to more general conceptual understandings about the phenomena. Finally, we align with Sandoval and Millwood's (2005) contention that the quality of the argument being presented by students as part of their inquiry investigations can be assessed. This assessment can be made based specifically on their claim and evidence or holistically as a collective statement.

Data Sources, Evidence, Objects, or Materials

Data for this study was derived from the piloting of the CyberPad application in two fourth grade elementary classrooms, one in each of two schools: NCA and LEA. NCA had 11 male and 12 female students, while LEA had 15 male and 11 female students (N=49 for both classrooms). Demographic data for the schools as a whole can be seen in Table 1. The same five units of instruction on electricity covered in both classrooms, though they were covered in four consecutive days in one school (NCA) and five days in the other (LEA). Five major topics in electricity were covered during the week:

1. Simple circuits
2. Energy conversion
3. Insulators and Conductors
4. Series circuits *
5. Parallel circuits *

   * Covered in a single day at NCA

Table 1 – School Demographics (percentages)

School	White	African American	Hispanic	Asian	Biracial	Economically Disadvantaged	Limited English Proficiency
LEA	0.59	0.23	0.07	0.07	0.05	0.24	0.06
NCA	0.59	0.09	0.28	0.02	0.02	0.34	0.20

Each topic was paired with activities from the FOSS™ Energy and Electromagnetism kit curriculum (FOSS™ Project, 2008) in which students engaged in content and responded to prompts on the CyberPad. The content on the CyberPad plus the kit material were designed to provide a rich inquiry cycle by which students would have the opportunity to consider focus questions for investigation around a science concept, make predictions about science phenomena, engage in observation and data collection of scientific phenomena, and reflect on their observations relative to the initial focus question, science concepts introduced to them, and their predictions. Log data from the CyberPad was harvested and both closed-ended responses (e.g., picking from a menu of items) and open-ended writing were analyzed. Across the five topics, there were 32 prompts requiring closed-ended responses and 51 requiring open-ended responses. Of the open-ended response prompts, 22 were structured to encourage students to respond in the form of an argument. These argumentation prompts are the focus of this current study. Of these argument prompts, 10 were hypothetical while 12 were empirical. In addition, students took a pre/post knowledge test consisting of 10 multiple-choice items drawn from released test items from NAEP, PISA, and North Carolina 5th grade End of Grade Science tests. The items were chosen based on their alignment to the learning goals established in the FOSS curriculum materials for electricity and meant to measure core conceptual content knowledge in electricity.

Methods, Techniques, or Modes of Inquiry

The overall project, that this study was a part of, is structured as a design-based project to build cyberlearning tools for enhanced student learning in classroom-based elementary physical science (Collins, Joseph, & Bielaczyc, 2004; Penuel, Fishman, Haugan Cheng, & Sabelli, 2011). The specific goal for this study was to use a rigorous argumentation framework as a lens for analyzing the substance and quality of written student argumentation for the purpose of guiding the software development and informing professional development around the use of the CyberPad.

The analysis presented here asked a pair of related questions regarding whether CyberPad prompts were able to elicit the kind of argumentative response expected by the designers and whether students were able to improve in their ability to respond to these argumentation prompts over the course of the week. To answer these questions, each prompt expected to elicit an argument was identified by the designers as either hypothetical (AH) or empirical (AE). Hypothetical prompts occurred prior to the data collection/observation portion of the activity. Empirical prompts came after data collection/observation and encouraged students to refer back (and use) both observational data and science conceptual knowledge. Based on the design of the prompts, some were expected to only result in a claim being made, while with others there was an expectation that either evidence, or claim and evidence would be present in the response.

Prompts identified as AH or AE were coded by two coders independently after two rounds of training on a training set of data from another school. These coders then compared codes and resolved differences in their codes. A third coder acted as tie-breaker on a handful of problematic student responses. In addition to coding the structure of the student response, a score of 0 (non-responsive to prompt) to 5 (ideal argument structure) was given. This scoring rubric assessed the quality of the argument but not the scientific accuracy of the responses. Finally, when student responses were coded, the presence of a claim, evidence, both, or neither was coded. In addition, the basis or source of the claim/evidence was coded as seen in Table 2.

Table 2 – Source of reasoning for argumentation prompts

Hypothetical (AH)
- Basic, or un-tutored assumption of how things work (B)
- Experience in previous activity (X)
- Science fact from current lesson (F)
- Off topic (O)

Empirical (AE)
- Reference to observation (R)
- Reasoning (inductive) (I)
- Abstraction of concept (A)
- Off topic (O)

Results and/or Substantiated Conclusions or Warrants

Table 3 shows the results of the pre/post knowledge assessment for the two schools in the study. Both schools showed significant gains with medium effect sizes, though NCA had a slightly larger effect size.

Table 3 – Pre/Post Knowledge Test Results

Classroom	N	Post *M*(*SD*)	Pre *M*(*SD*)	M_{diff}	*t*(df)	*p*	*Cohen's d*
LEA	24	6.25 (2.09)	5.00 (1.93)	1.25	2.85(23)	0.01	0.33
NCA	20	6.33 (2.43)	4.75 (1.74)	1.58	3.69(19)	0.00	0.75

Table 4 provides a summary of tallies of student responses (actual) relative to the expected number to be elicited by the prompts. These expected values reflect which of the AH and AE prompts designers expected claim only, evidence only, or claim and evidence responses.

Table 4 – Expected and actual tallies of student responses to argumentation prompts

School	LEA		NCA	
Prompt Element				
AE	Expected	Actual	Expected	Actual
Claim Only	171	173	143	113
Evidence Only	0	25	0	32
Claim + Evidence	97	23	82	38
No Claim or Evidence	0	6	0	2
No Answer	0	16	0	21
Unproductive	0	23	0	19

AH				
Claim Only	149	150	147	112
Evidence Only	24	18	24	22
Claim + Evidence	25	15	18	12
No Claim or Evidence	0	6	0	2
No Answer	0	2	0	1
Unproductive	0	7	0	6

A closer look was then made at a specific pair of hypothetical prompts that were expected to only elicit claims-one from the first day's activity where simple circuits were explored (P1_6):

What is needed to make a light bulb light?

and one from the fifth day's activity where parallel circuits were explored (P5_1):

How can you light two bulbs brightly with just one D-cell? Answer this question based on what you know about circuits.

The student responses, as coded, are seen in Table 5.

Table 5. Student responses to the Prompts P1_6 and P5_1, reported by source of response (N=49). Refer to Table 2 for the source definitions.

Source	**Prompt P1_6**	**P5_1**
R*	0	1
B	2	4
X	0	37
F	41	1
O	6	6

* Note: Reference to an observation (R) was not expected in a hypothetical prompt

A one-tailed t-test of the hypothesis that the quality of prompt P5_1 would be greater than prompt P1_6 indicated that this was indeed the case (P1_6, M=2.45; P5_1, M=2.72; $t(48)=1.93$, $p > 0.029$).

A similar comparison was made between two empirically based prompts-P2_9 from an activity on further explorations of simple circuits and energy conversion:

To answer the starting question, explain how the switch turns the electricity on and off when you open it and close it. Here's a hint: remember that a complete or closed circuit is necessary to run the motor.

and P4_11, from an activity on series circuits:

How can you make the two bulbs in a series circuit brighter? Answer this question based on what you already know about circuits and explain why you think your solution would make the bulbs brighter.

The student responses, as coded, are seen in Table 6.

Table 6. Student responses to the Prompts P2_9 and P4_11, reported by source of response (N=49).

Source	Prompt P2_9	P4_11
A	4	17
I	3	11
R	33	13
O	9	8

A one-tailed t-test of the hypothesis that the quality of prompt P4_11 would be greater than prompt P2_9 indicated that this was not the case (P2_9, M=2.58; P4_11, M=2.50; $t(48)=0.52$ $p > 0.302$).

The results of the pre/post knowledge test indicates that students demonstrated learning around conceptual knowledge that is traditionally taught and tested in elementary science. However, the next generation science standards framework (NRC, 2011) and other work in reform-based science instruction say that developing competency in practices such as argumentation has a separate set of instructional challenges. The results of this study suggest that with argumentation prompts for which just a claim was expected, students were largely successful in providing this (though LEA was slightly better than NCA). However on prompts where evidence was also expected, students at both schools struggled more with this, especially for empirical argumentation prompts.

Drilling down into what students wrote on specific prompts reinforces some of the differences seen in responses for AH and AE prompts. For AH prompt P1_6, students were largely successful at integrating content knowledge from the current lesson into their claim statements. Similarly, for prompt P5_1, on the fifth day of the activity, students were able to bring content knowledge from prior days to bear on their claim response. In addition, over the course of the week, students seem to improve in the quality of their responses. However, for the AE prompts, a smaller percentage of students were able to make use of data from their labs to provide evidence for their claims. This was true for both P2_9 and P4_11. In fact, for P4_11 a more conceptually complex lab on series circuits students, students had overall weaker responses than for the earlier activity associated with P2_9. That said, a minority of students responding to P4_11 were able to formulate inductive responses that moved from the specific observations in the lab to more general statements about the science concept.

Scientific or Scholarly Significance of the Study

This study is demonstrative of the type of analytic work that can be done by mining student work in order to inform the development of cyberlearning systems for elementary science education. In and of itself, it demonstrated the difficulties students have in bringing evidence to bear in formulating scientific arguments. It also points up the challenge that the Leonardo Project and similar cyberlearning projects have in scaffolding the development of effective epistemic frames for learning and communication of findings in the form of strong arguments.

References

Berland, L. K., & Reiser, B. J. (2008). Making sense of argumentation and explanation. *Science Education, 93*, 26-55.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the learning sciences, 13*(1), 15-42.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education, 84*(3), 287-312.

FOSS Project. (2008). Welcome to FOSSWeb Retrieved July 1, 2009

Leonardo Project. (2011). Homepage Retrieved July 16, 2012, from http://www.intellimedia.ncsu.edu/leonardo.html

McNeill, K. L. (2011). Elementary students' views of explanation, argumentation, and evidence, and their abilities to construct arguments over the school year. *Journal of Research in Science Teaching, 48*(7), 793-823. doi: 10.1002/tea.20430

McNeill, K. L., & Krajcik, J. S. (2008). Scientific Explanations: Characterizing and Evaluating the Effects of Teachers‟ Instructional Practices on Student Learning. *Journal of Research in Science Teaching, 45*(1), 53-78.

NRC, National Research Council. (2011). A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. Washington, DC: The National Academies.

Penuel, W. R., Fishman, B. J., Haugan Cheng, B., & Sabelli, N. (2011). Organizing Research and Development at the Intersection of Learning, Implementation, and Design. *Educational Researcher, 40*(7), 331-337. doi: 10.3102/0013189x11421826

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., & Millwood, K. A. (2005). The Quality of Students' Use of Evidence in Written Scientific Explanations. *Cognition and instruction, 23*(1), 23-55. doi: 10.1207/s1532690xci2301_2

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

Varelas, M., Pappas, C. C., Kane, J. M., Arsenault, A., Hankes, J., & Cowan, B. M. (2008). Urban primary-grade children think and talk science: Curricular and instructional practices that nurture participation and argumentation. *Science Education, 92*(1), 65-95. doi: 10.1002/sce.20232