Argumentation and Epistemic Criteria: Investigating Learners' Reasons for Reasons

Richard Duschl[1] and Kirsten Ellenbogen[2]

ABSTRACT

The dominant practices in science are not discovery and justification but rather the broadening and deepening of explanations via theory development and conceptual modification. New tools, theories, and technologies are fundamentally changing the methods and inquiry practices of science. In turn, the forms of evidence and criteria for what counts as evidence change, too. The implication is the importance of including and understanding the epistemic and social practices embedded in the science learning goals 'critique and communication of scientific ideas and information'. The study examines 11-year old students' argumentation patterns and epistemic reasoning. The focus is on students' arguments and choices as they pertain to reasons and evidence given to support claims about data and measurement. Results suggest that students are able to evaluate the quality of arguments made by classmates and can do so employing a number of different reasons or perspectives. An understanding of how students choose and use evidence is important for understanding how to coordinate formative assessments and focus teacher's feedback on students' epistemic reasoning.

KEYWORDS: Argumentation, Computer Supported Learning, Scientific Reasoning, Philosophy of Science.

Introduction

The last 50 years have been a parade of new scientific methods, ideas and beliefs; Pluto is no longer considered a planet! The practices and images of science today are so very different and this is largely due to the advancements in methods and instrumentation for observing, measuring, analyzing and modeling. Certainly the fundamental frameworks of scientists working from material evidence to propose explanations and mechanisms are generally the same. New tools, theories, and technologies though have fundamentally changed both the guiding frameworks and the inquiry practices of science (Duschl & Grandy, 2008; NRC, 2007a). In turn, the forms of evidence and criteria for what counts as evidence have had to change, too.

The contemporary understanding of the nature of science (c.f., Giere, 1988; Longino, 2002; Godfrey-Smith, 2003; Soloman, 2008) holds that the majority of scientists' engagement is not individual efforts toward establishing final theory acceptance, but rather communities of scientists striving for improvement and refinement of a theory. What occurs in science is not predominantly the context of discovery or the context of justification but the contexts of theory development, of conceptual modification. Thagard (2007) posits that coherence, truth, and explanatory coherence of scientific explanations are achieved through the complementary process in which theories broaden and deepen over time by accounting for new facts and providing explanations of why the theory works. Recent research reviews (NRC, 2007a; Duschl, 2008; Duschl & Grandy, 2008; Ford & Forman, 2006; Lehrer & Schauble, 2006b) and research studies on science learning (Ford, 2008a; Lehrer, Schauble, & Lucas, 2008; Smith, Wiser, Anderson & Krajcik, 2006) maintain that the same broadening and deepening practices based on a set of improving and refining tenets within a community of investigators ought to hold for teaching and learning practices in science learning environments. For scholars in the emerging domains of the learning sciences and science studies the emphasis is on 'science as practice' and in particular the epistemic and social practices embedded in the critique and communication of scientific ideas and information.

The research reported here is part of a broader research program that seeks to better understand how to mediate students' argumentation discourse (Duschl, 2008; Erduran & Jimenez-Aleixandre, 2008; Duschl & Osborne, 2002). Here we report results of a study that examines students' reasoning as it pertains to reasons and evidence given to support claims about data and measurement; i.e., what counts as evidence for the critique of data. An understanding of how students choose and use evidence is important to understanding how to coordinate and focus teacher's feedback on epistemic reasoning and learners' practices with building and refining models.

Review of Literature

Wellington and Osborne (2001) argue that if science and scientists are epistemically priviledged, then it is a major

---

[1] Department of Curriculum and Instruction, Penn State University, University Park, PA USA.
[2] Science Museum of Minnesota, St. Paul, MN, USA.

shortcoming of our educational programs that we offer little to justify the current lack of focus on epistemic practices in classrooms. Wellington and Osborne are speaking to the misplace priorities we find in most science curriculum. That is, the dominant focus is teaching what we know. How we came to know and why we believe what we know are marginalized aspects of science learning. Successful science education depends on students' involvement in forms of communication and reasoning that models the discourse that occurs in scientific communities (e.g., Herrenkohl, Palincsar, DeWaters, Kawasaki, 1999; Roseberry, Warren & Conant, 1992; Schauble, Glaser, Duschl, Schulze, & John, 1995).

A critical step forward is engaging learners in examining the relationships between evidence and explanation. A dominant dynamic in scientific communication is the dialectical practices that occur between evidence and explanation or, more generally, between observation and theory. Increasingly when examined over the last 100 years, these dialectical exchanges involve or depend on the tools and instruments scientists employ (Zammito, 2004). Ackerman (1985) refers to such exchange practices as the 'data texts' of science and warns that the conversations among contemporary scientists about measurement, observations, data, evidence, models, and explanations is of a kind that is quite foreign from the conversations found in the general population and, we might add, in science classrooms (Duschl, 2008; Driver, Newton & Osborne, 2000).

Pickering (1995) referred to this conflation of conversations when describing experiments in high-energy physics as the 'mangle of practice'. Hacking (1988) writes that it is the richness, complexity and variety of scientific life that has brought about a focus on 'science as practice.' For Pickering, scientific inquiry during its planning and implementation stages is a patchy and fragmented set of processes mobilized around resources. Planning is the contingent and creative designation of goals. Implementation for Pickering (1989) has three elements:

- a "material procedure" which involves setting up, running and monitoring an apparatus;
- an "instrumental model," which conceives how the apparatus should function; and
- a "phenomenal model," which "endows experimental findings within meaning and significance … a conceptual understanding of whatever aspect of the phenomenal world is under investigation.

The "hard work" of science comes in trying to make all these work together (Zammito, 2004; pp. 226-227).

A critical step forward toward changing the 'what we know' instructional condition is to engage learners in examining the relationships between evidence and explanation. As stated above, a dominant dynamic in scientific communication is the data text discourse that occurs between evidence and explanation. This dialectic typically takes the form of arguments and is central to science-in-the-making activities.

Duschl (2003; 2008) has developed a set of design principles that coordinate conceptual, epistemic and social learning goals. The pedagogical model is the Evidence-Explanation (E-E) continuum depicted schematically in Figure 1.

The E-E continuum (Duschl, 2003) has its roots in perspectives from science studies and connects to cognitive and psychological views of learning. The appeal to adopting the E-E continuum as a pedagogical framework for science education is that it helps work out the details of the epistemic discourse processes. It does so by formatting into the instructional sequence select junctures of reasoning, e.g., *data texts transformations*. At each of these junctures or transformations, instruction pauses to allow students to make and report judgments. The three critical transformations or judgments in the E-E continuum include:

a) Selecting or generating data to become evidence.
b) Using evidence to ascertain patterns of evidence and models.
c) Employing the models and patterns to propose explanations.

The results reported in this paper focus on the first and second transformations; i.e., the reasons given for what counts as evidence and patterns in evidence. Yet another important judgment is deciding what data to obtain (Lehrer & Schauble, 2006a; Petrosino, Lehrer & Schauble, 2003). The development of children's consideration for measurement and data prior to starting the E-E continuum is critically important, too.

Research Context and Methods

The present study examines two classes of 11-year old students' argumentation patterns and epistemic reasoning. The school setting was a K-12 international school for American children located in a major metropolitan city. One of many located around the world, the school provides an American education for families living abroad. The school is fee paying and the students come from a mix of middle and high SES families. The school was physically divided into three sections
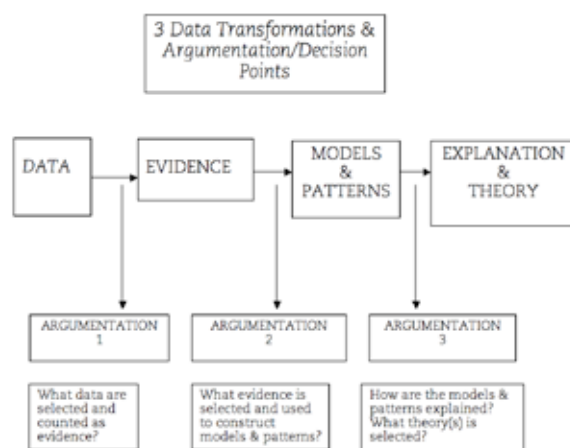

Figure 1. 3 Data Transformations and Decision Making Points.

Image 1. *Knowledge Forum* Note with 'Theory Building' prompts.

— the lower school grades K-5 (ages 5-10), the middle school grades 6-8 (ages 11-13), and the upper school grades 9-12 (ages 14-18). The students for this study were 6th graders thus part of the middle school. The middle school had two instructional teams made up of four teachers — English, Social Studies, Science and Math. In addition, the middle school had a full-time teacher assigned to the dedicated space for computer instruction. For this study, we worked with each of the science teachers and the computer teacher. Teachers at the school are typically American educated but not necessarily in possession of a teaching certificate.

The science subject context is a two-week unit on "Exercise for a Healthy Heart" developed by the American Heart Association (AHA). Students participated in several labs that require taking a pulse and gathering pulse data. There were three conditions — at-rest, change of pace, and change of weight — that students completed in three sequential lessons in the science classroom. After all the data gathering of pulse rates was completed students moved to the middle school computer lab. Each student was provided a computer station. The "Knowledge Forum" (KF) computer software program was loaded onto each of the computers and the classroom server. KF was formerly known as CSILE (Computer Supported Intensive Learning Environments) and was designed with scaffolding tools and prompts to support learners' knowledge building and knowledge revision activities (Scardemalia, Bereiter & Lamon, 1994). In a 'chat room' type of environment, pupils can post notes, read classmates notes and build-on and make collections of notes (see Image 1).

For the present study, scaffold tools located in KF were adapted to help guide and mediate pupil's participation in 'speaking together' sessions. One adaption was to overlay the KF discourse environment with the Project SEPIA 'assessment conversation' framework (Duschl, 2003; Duschl & Gitomer, 1997). Speaking together opportunities when properly planned and managed are occasions for making thinking visible. In Project SEPIA the 'speaking together' occasions take place during instructional episodes designated assessment conversations (AC). The AC is a teaching model for facilitating and mediating discourse practices and it has three general stages:

1. *Receive ideas.* Students 'show what they know' by producing detailed writing, drawings, and storyboards.
2. *Recognize ideas.* The diversity of ideas received is made public. Students' ideas and the data from investigations are then discussed against a set of scientific criteria and evidence. The teacher selects the student work that will reveal the critical differences in student representations and reasoning.
3. *Use the ideas.* A class discussion takes place that examines and debates student's representations, reasons, and thinking. The teacher's role is to pose questions and facilitate discussion that results in a consensus view(s).

A second adaptation to KF was to change the writing scaffolding prompts in the software to argumentation prompts. The argumentation prompts focused students on considering what counts as good and accurate data; the Argumentation 1 and 2 transformation points shown in Figure 1. Two probes of inquiry were used — one that asked students to consider the collection of data and one that engaged them in considering the analysis of data.

Following the completion of pulse data gathering investigations, students were presented with KF tasks to pursue students' arguments concerning the collection of data. In the first activity students were directed to write and post a 'note' agreeing or disagreeing with each lettered statement below. The KF note window prompted students to supply a reason and to cite evidence to support an agree or a disagree position. The four data statements presented to students on the KF screen as notes are:

A. It matters where you take a pulse.
B. It matters how long you take a pulse (6, 10, 15 or 60 seconds).
C. It matters when you begin to take a pulse after exercising.
D. It matters who takes a pulse.

Upon completion of the write/post-a-note activity, we found that of the four statements, only the first two statements A and B produced a diversity of responses suitable for an assessment conversation on coordinating and conducting argumentation discourse activities. For this research report, our focus is only on Statement B - "It matters how long you take a pulse." Examples of students posted notes (as written and submitted) for Statement B include:

• **Note 1**   It doesn't really matter for how long you take the pulse, just as long as you multiply the pulse by the correct number of seconds.
• **Note 2**   I think it does matter how long you take the pulse. The only true way to get the heart rate per minute is to take the pulse rate for a minute.
• **Note 3**   Yes, I think time does matter in taking a pulse. Such as being active. You stop being active and take your pulse. The longer you take, you heart beat will slow down to it normal rate.

B-A&D-1   5 votes

Statement B: It matters how long you take a pulse *Disagree* **Reason Why** If you are working out and you want to take your pulse you have to do it right after you stop your activity. *Agree* **Reason Why** If you are taking your resting heart rate it doesn't matter how long you take it for because it will be consistant (sic).

Reasons for Selecting B-A&D-1 as a Good and Accurate Statement:
— You could really understand what they were talking about and saying.
— They explained themselves and I think they were right.
— They have a good argument and because if you only take a pulse for two seconds and times it by forty it won't be accurate unless your lucky.
— One of these explains agree really well with detail and another explains both pretty well.
— They are accurate and they prove the point well.

Figure 2. Reasons given for highly selected Best Statement B-A&D-1.

- **Note 4**   I think taking the pulse for 60 seconds is more accurate because if you take a pulse for 30 secs., 15 sec., or 10 sec. you have to multiply and you can make a mistake.
- **Note 7**   Yes because there heart bets different.
- **Note 8** I don't think it matters because it is all the same because you keep the same constant pulse if you take it in that period of time.
- **Note 9**   I think it matters a lot because the more we move the more blood we use and the faster we go.
- **Note 10**   I do not think it matters because, you only have one constant pulse rate.
- **Note 24**   I actually think that taking people's pulse for 1 whole minute is the most accurate. I think that because You don't have to times the number by another number. If you take someone's pulse for 6 seconds, and multiply wrong or something like that, your whole pulse would be wrong!!!!!
- **Note 25**   I think it would be more acurate to take a pulse for 15 sec. You can mess up if u take a pulse with your thumb, lose count, or if the other person gets impatient.

From the total pool of students' responses, the researchers selected a sample of 15 responses from the two classes (eight agree, six disagree and one agree & disagree) as responses that reflected the diversity of student's ideas and positions. These 15 responses were posted on KF as 'notes' on a new view. A 'view' is the label used in KF to describe a new forum for student discussion. The next class session when students arrived to the computer classroom the 15 notes were on the computer desktop. The students were directed to open, read

B-A-3   8 votes

Statement B: It matters how long you take a pulse *Agree* **Reason Why** Because if you are recording the pulse after a certain amount of exercise because the blood pressure will start to go down after a certain amount of time, therefore making the number of beats per minute. It would be most accurate if you took the pulse for ten seconds multiplying it by six to find the number of beats per minute.

Reasons for Selecting B-A-3 as a Good and Accurate Statement:
— Give good support.
— It has a good point and backs up its point very well.
— It is clear and it explain the argument well (ie. the blood pressure will start to go down after a certain amount of time).
— They have a good argument and because if you only take a pulse for two seconds and times it by 40 it won't be accurate unless your lucky.
— They make their point. They brought new ideas and explained them. They also took old ideas but explained them better than the others.
— I agree with them and they prove they're point well. They're point is that if one takes one's pulse for 3 secs. you would probably not find an accurate heart beat.
— They say that it does matter and then explane why it matters.
— They are all making a different point and that makes more clear to understand the answers.

Figure 3. Reasons given for highly selected Best Statement B-A-3.

and respond to each of the notes. For each note, students were directed to use the KF scaffold tool: 1) to either **agree** or **disagree** with the statement; 2) to provide a reason for this decision, and 3) to cite evidence to support the reason.

The most frequently selected notes with accompanying arguments (reasons and evidence) can be found in Figures 2, 3, 4, and 5.

The codes B-A-8 refer to statement B, Agree, and statement 8. The **boldface** words signify scaffold tool buttons we the researchers embedded into the KF environment. Students were instructed to use these tools as they wrote the agree/disagree statements. By design, as stated above, a diversity of student responses were selected for the purpose of studying what criteria students would use to assess and evaluate the position made by students in their statements.

Table 1 presents the outcome of the task. Three statements were judged by students to be the best. Namely, B-A-3 with eight votes; B-A-8 with six votes and B-A&D –1 with five votes. While the intention of the researchers was that students would make only one selection, many students followed the make-a-collection instructions and selected two or more best statements.

Results/Points of View
Consensus opinions about the best arguments did emerge

B-A-8   6 Votes

Agree
Reason Why Because if you take a pulse for too long, the person you are taking the pulse of might calm down and it would beat slower.
We took pulses in science class and after resting the pulse was slower, and after doing step-aerobics it was faster.

Reasons for Selecting B-A-8 as a Good and Accurate Statement:
— They explained themselves and I think they were right.
— One of these explains agree really well with detail and another explains both pretty well.
— No reason given.
— They say that it does matter and then explane why it matters.
— They are all making a different point and that makes more clear to understand the answers.

Figure 4. Reasons given for highly selected Best Statement B-A-8.

B-D-5   3 Votes

Statement B: It matters how long you take a pulse
Disagree
Reason Why It does not matter if you are resting but if you have been exercising and want to find the pulse you must take it as soon as possible or it will slow down. Evidence After I had been exercising I had a glass of water and then took my pulse but it was the same as my resting heart rate.

Reasons for Selecting B-D-5 as a Good and Accurate Statement:
— They are accurate and they prove the point well.
— They make their point. They brought new ideas and explained them. They also took old ideas but explained them better than the others.
— They are all making a different point and that makes more clear to understand the answers.

Figure 5. Alternative best statement selected along with highest vote best statement.

---

from the KF tasks. The data show that 12 students — the consensus group — choose at least one of the three top selected statements. Looking across the three statements and focusing on those students who opted for more than one choice, we see a further consensus among the students. That is, of the students making more than one choice many opted for one or the other of the top three selected statements. This is shown in brackets on the end of reason statements (e.g., [B-D-5].

Following the lead of these 12 students, and with an eye toward establishing the context for the next assessment conversation on the issue of what counts as a good reason, the other selections of the consensus group were examined. Three of the consensus group students selected statement B-D-5, see Figure 5. This represents the only disagree statement. A careful read of the statement shows a contradiction by the student writer in that the narrative of the reason reflects agreement with statement B. This statement is also consistent with the other top selected statements in that it addresses the problem of the heart rate slowing down or calming down with time. Finally, it is the only statement among the consensus group that employed the evidence scaffolding tool. This is significant since a goal of the instruction is to model, scaffold and coach students into the practice of making arguments that site evidence.

We also found it interesting to look at the selection of statements from the student who choose seven and used the same reason for each of the selections - THEY HAVE A GOOD ARGUMENT AND BECAUSE IF YOU ONLY TAKE A PULSE FOR TWO SECONDS AND TIMES IT BY FORTY IT WON'T BE ACCURATE UNLESS YOUR LUCKY. The seven selections are found in Table 1. A review of the seven statements reveals a theme that is different from the other eight statements. That is, a concern for the magnification of measurement error when multiplying short durations of data collection (e.g., two or three seconds) to infer beats/minute. One can imagine how a subsequent assessment conversation can focus students' attention on the role of measurement, the accuracy of measurement and the ways in which the spread of measurements can influence outcomes in an investigation. Such interrogation of data has been shown to be a critically important step toward students' engaging in a deepening and broadening of explanatory models and mechanisms (Lehrer, Schauble & Lucas, 2008; Petrosino, et al., 2003; Lehrer & Romberg, 1996). What this program of research has established, and what we contend an exploration of reasons about reasons can reveal, is the confusion students typically make between the attributes and the properties of materials and organisms.

For the data collection task, the research data show that the majority of students are able to evaluate the quality of arguments. For 12 students a consensus of opinion emerged on what counts as accurate data for an exercising heart rate. We also see an alternative perspective on counting errors emerged with various suggestions on how to deal with counting errors. For example, recommendations to take the pulse for a full 60 seconds may work for resting heart rate but this strategy will not work for an exercising heart rate. With the passage of 60 seconds the pulse rate will slow down. Only some of the students see this situation and take it into account for the evaluation of the statements.

From Table 1 we can see that there were six statements that did not receive any votes.

These results suggest that students are able to evaluate the quality of arguments made by classmates and can do so employing a number of different reasons or perspectives. Disturbingly, the data also reveal that students are not providing evidence for the reasons selected. Students did not employ the KF evidence scaffold tool. This suggests there is confusion among students about the difference between reasons and evidence or the need to support reasons with evidence.

Table 1. Statement Selections and Reasons for Selection

Statement B: It matters how long you take a pulse.

Statement B	Number of Students Selecting Statement	Reasons for Selecting Statement a
B-A&D-1	XXXXX 5	— You could really understand what they were talking about and saying; — They explained themselves and I think they were right(2); — They have a good argument and because if you only take a pulse for two seconds and times it by forty it won't be accurate unless your lucky (7); — One of these explains agree really well with detail and another explains both pretty well (2); — They are accurate and they prove the point well (2).
B-A-1	X 1	— They have a good argument and because if you only take a pulse for two seconds and times it by forty it won't be accurate unless your lucky (7).
B-A-2	XX 2	— They have a good argument and because if you only take a pulse for two seconds and times it by forty it won't be accurate unless your lucky (7); — No reason given (2).
B-A-3	XXXXXXXX 8	— Give good support(2); — It has a good point and backs up its point very well; — It is clear and it explain the argument well (ie. the blood pressure will start to go down after a certain amount of time; — They have a good argument and because if you only take a pulse for two seconds and times it by forty it won't be accurate unless your lucky (7); — they make their point. They brought new ideas and explained them. They also took old ideas but explained them better than the others. (2); — I agree with them and they prove they're point well. They're point is that if one takes one's pulse for 3 secs. you would probably not find an accurate heart beat(2); — They say that it does matter and then explane why it matters (2); — They are all making a different point and that makes more clear to understand the answers (3);
B-A-4	XX 2	— They have a good argument and because if you only take a pulse for two seconds and times it by forty it won't be accurate unless your lucky (7); — I agree with them and they prove they're point well. They're point is that if one takes one's pulse for 3 secs. you would probably not find an accurate heart beat(2).
B-A-5	0	
B-A-6	XX 2	— It says all you need to know and it uses a good example; — They have a good argument and because if you only take a pulse for two seconds and times it by forty it won't be accurate unless your lucky (7)
B-A-7	XX 2	— Give good support(2); — They have a good argument and because if you only take a pulse for two seconds and times it by forty it won't be accurate unless your lucky (7)
B-A-8	XXXXXX 6	— They explained themselves and I think they were right (2); — One of these explains agree really well with detail and another explains both pretty well (2); — No reason given (2); — They say that it does matter and then explane why it matters (2); — They are all making a different point and that makes more clear to understand the answers (3).
B-D-1	0	
B-D-2	0	
B-D-3	0	
B-D-4	0	
B-D-5	XXX 3	— They are accurate and they prove the point well (2); — They make their point. They brought new ideas and explained them. They also took old ideas but explained them better than the others. (2); — They are all making a different point and that makes more clear to understand the answers (3).
B-D-6	0	

a- Number in ( ) indicates the number of times a student used the 'reason statement' to justify selecting a Agree/Disagree statement. For example, one student used the following statement seven times: THEY HAVE A GOOD ARGUMENT AND BECAUSE IF YOU ONLY TAKE A PULSE FOR TWO SECONDS AND TIMES IT BY FORTY IT WON'T BE ACCURATE UNLESS YOUR LUCKY.

Educational/Scientific Importance

An understanding of how students choose and use evidence is a goal of this research study. Making thinking visible about choosing and using evidence is important to understand so that we can help to coordinate and focus teacher's feedback on epistemic reasoning. Additionally, choosing and using evidence are core practices learners' need to use when building and refining models.

Students' using the analysis of the heart beat patterns and graphical representations to comment on issues about good and accurate data was successful to a point. We found that the development of the ability to distinguish reasons from evidence is an instructional situation that will need careful mediation from the teacher and from knowledgeable peers. It would seem that more attention needs to be given to the interrogation and modeling of the pulse rate data. Lehrer and Schauble (2006b) have reported that getting students to engage in resemblance representation tasks is an entrée to modeling. Modeling, they maintain, is then sustained and extended by using resemblances to construct representational forms that afford quantification and investigation of relations among quantities (Lehrer & Schauble, 2006a; Lehrer et al., 2008).

There is also the possibility that the inability to distinguish reasons from evidence is related to confusions surrounding attributes (e.g., blood flow) and properties (e.g., pulse rate). The significance of this possibility is that domain-general rules for what counts as a reason, evidence, warrants and backings so not work effectively across different investigative contexts (Lehrer et al., 2008). Rather, such 'what counts' rules will need to be sensitive to both the conceptual and the epistemological domain under investigation. The development of learners' abilities to talk about measurement, observations, evidence, patterns, and modeling will need frequent and insightful mediation from teachers and more knowledgeable peers.

This research study has raised questions and issues about the design of science learning environments that seek to promote scientific discourse practices. We see the following assertions and issues as relevant to the enterprise:

- The ability for students and teachers to reconfigure information allows post hoc discussions about evidence, methods of data collection, reasoning, and explanations. Computer supported discourse like KF allows the pace of argumentation to slow down and thus enable teachers and students to 'take stock' of the information and ideas emerging from scientific inquiries.
- Our interventions show that ideas and information emerging from small group and whole class activities can be captured by a computer server and used effectively to inform assessment decision making and the mediation of students epistemic reasoning.
- One of the main advantages of the computer supported environments like KF is that the visibility of student thinking is available for inspection and re-inspection by the teacher and students alike. Employing the KF platform in conjunction with other artifacts (e.g., notebooks, journals) makes possible studying the fate or outcome of a line of reasoning. We see very important implications for understanding the developmental corridors for learning epistemic practices of science. Knowledge Forum, in many ways, predates the structures and philosophies of "Web 2.0." Web 2.0 is exemplified by wikis, blogs, and user-driven websites that value user-generated content. These environments include creating archives as collective memory, visualizing cognitive processes, and uniting groups that are geographically distributed (Andriessen et al., 2003; Zimmerman, 2005). Drawing on these theories about the nature of CSCL learning, researchers have reported on a number of structured online systems developed to facilitate individual and group science learning by emphasizing the formation of argumentative discourse through group conversations and writing projects (Bell, 2000; DeVries, Lund, & Baker, 2002; Scardamalia, 2002).
- Our study suggests that the ability to manipulate the metacognitive prompts (i.e., scaffold tool bar) is a powerful way to both engage and develop epistemic reasoning. When students become aware of their own learning processes, they gain much richer understandings of the content of their learning, and become better, more empowered learners (e.g. Baird, 1986). Metacognitive prompts can focus on students' self-reflections of their interactions and roles, their self-perceptions of the learning tasks, and their learning strategies. Students can also be asked to explain and elaborate on any other strategies that they promoted their group and individual activities, learning, and task completion.

Bibliography

Ackerman, R.J., *Data, instruments, and theory: A dialectical approach to understanding science.* Princeton, NJ: Princeton University Press, 1985.

Andreissen, J., Bake, M., & Suthers, D., *Arguing to Learn: Confronting Cognitions in Computer-Supported Collaborative Learning Environments.* Dordrecht/Boston/London: Kluwer Academic Publishers, 2003.

Baird, J.R., Improving learning through enhanced metacognition: A classroom study, *European Journal of Science Education*, **8**(3), 263-282, 1986.

Bell, P., Scientific arguments as learning artifacts: Designing for learning from the web with KIE, *International Journal of Science Education*, **22**(8), 797-817, 2000.

De Vries, E., Lund, K., & Baker, M., Computer-mediated epistemic dialogue: explanation and argumentation as vehicles for understanding scientific notions, *The Journal of the Learning Sciences*, **11**(1), 63-103, 2002.

Driver, R., Newton, P., and Osborne, J., Establishing the norms

of scientific argumentation in classrooms, *Science Education*, **84**(3), 287-312, 2000.

Duschl, R.A., Assessment of inquiry. In: J.M. Atkin & J.E. Coffey (eds.), *Everyday assessment in the science classroom* (pp. 41-59). Washington, DC: National Science Teaching Association Press, 2003.

Duschl, R., Science education in 3 part harmony: Balancing conceptual, epistemic and social learning goals. In: J. Green, A. Luke, & G. Kelly (eds.), *Review of Research in Education*, vol. 32 (pp. 268-291), Washington, DC: AERA, 2008.

Duschl, R.A. & Gitomer, D., Strategies and challenges to changing the focus of assessment and instruction in science classrooms, *Educational Assessment*, **4**(1), 337-73,1997.

Duschl, R. & Grandy, R. (eds.), *Teaching Scientific Inquiry: Recommendations for Research and Implementation*. Rotterdam, Netherlands: Sense Publishers, 2008.

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

Erduran, S. & Jimenez-Aleixandre, M.P. (eds.)., *Argumentation in science education: Perspectives from classroom-based research*. Dordrecht, Netherlands: Springer, 2007.

Ford, M., Disciplinary authority and accountability in scientific practice and learning, *Science Education*, **92**(3), 404-423, 2008a.

Ford, M., 'Grasp of practice' as a reasoning resource for inquiry and nature of science understanding, *Science & Education*, **17**, 147-177, 2008b.

Ford, M.J. & Forman, E.A., Redefining disciplinary learning in classroom contexts. *Review of Research in Education*, **30**, 1-32, 2006.

Giere, R., *Explaining science: A cognitive approach*. Chicago: University of Chicago Press, 1988.

Godfrey-Smith, P., *Theory and reality*. Chicago: University of Chicago Press, 2003.

Hacking, I., Philosophers of Experiment. In: A. Fine & J. Leplin (eds.), *PSA 1988*, vol. 2 (pp. 147-156). East Lansing, MI: Philosophy of Science Association, 1988.

Herrenkohl, L.R., Palinscar, A.S., DeWater, L.S., & Kawasaki, K., Developing scientific communities in classrooms: a sociocognitive approach, *Journal of the Learning Sciences*, **8**, 451-493, 1999.

Lehrer, R. & Romberg, R., Exploring children's data modeling, *Cognition and Instruction*, **14**(1), 69-108, 1996.

Lehrer, R., & Schauble, L., Scientific thinking and science literacy. In: W. Damon, R. Lehrer, K.A. Renninger and I.E. Sigel (eds.), *Handbook of child psychology*, 6th edition, vol. 4, Child psychology in practice. Hoboken, NJ: Wiley, 2006a.

Lehrer, R. & Schauble, L., Cultivating model-based reasoning in science education. In: K. Sawyer (ed.), *The Cambridge handbook of the learning sciences*. New York: Cambridge University Press, 2006b.

Lehrer, R., Schauble, L., & Lucas, D., Supporting development of the epistemology of inquiry, *Cognitive Development*, **23**, 512-529, 2008.

Longino, H., *The fate of knowledge*. Princeton, NJ: Princeton University Press, 2002.

National Research Council., *Rising above the gathering storm: Energizing and employing America for a brighter economic future*. Washington, DC: National Academy Press, 2007a.

National Research Council, *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academy Press, 2007b.

Petrosino, A., Lehrer, R., & Schauble, L., Structuring error and experimental variation as Distribution in the fourth grade, *Mathematical Thinking and Learning*, **5**(2/3), 131-156, 2003.

Pickering, A., Living in the material world: On realism and experimental practice. In: D. Gooding, T. Pinch, & S. Schaffer (eds.), *The uses of experiment: Studies in the natural sciences* (pp. 275-298). Cambridge, UK: Cambridge University Press, 1989.

Pickering, A., *The mangle of practice: Time, agency and science*. Chicago: University of Chicago Press, 1995.

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

Scardamalia, M., Collective cognitive responsibility for the advancement of knowledge. In: B. Smith & C Bereiter (eds.) *Liberal Education in a Knowledge Society*. Chicago: Open Court Publishing, 2002.

Scardamalia, M., Bereiter, C., & Lamaon, M., The CSILE Project: Trying to bring the classroom into world 3. In: K. McGilly (ed.), *Classroom Lessons: Integrating Cognitive Theory and Classroom Practice* (pp. 201-228). Cambridge, MA: MIT Press, 1994.

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

Smith, C., Wiser, M., Anderson, C., &Krajcik, J., *Implications of research on children's Learning for assessment: Matter and atomic molecular theory. Measurement: Interdisciplinary Research and Perspectives*, vol. 4 (pp. 11-98). Mahwah, NJ: Lawrence Erlbaum, 2006.

Solomon, M., Social Epistemology of Science. In: R. Duschl & R. Grandy (eds.), *Teaching Scientific Inquiry: Recommendations for Research and Implementation*. Rotterdam, Netherlands: Sense Publishers2008.

Thagard, P., Coherence, Truth, and the development of scientific knowledge, *Philosophy of Science*, **74**(1), 28-47, 2007.

Wellington, J. & Osborne, J., *Language and literacy in science education*. Buckingham, UK: Open University Press, 2001.

Zammito, J.H., *A nice derangement of epistemes: Post-positivism in the study of science From Quine to Latour*. Chicago: University of Chicago Press, 2004.

Zimmerman, H.T., *Adolescent argumentation in online spaces*. Presentation at First Congress of the International Society for Cultural and Activity Research (ISCAR), Sevilla, Spain, September 2005.