Aggregating Students’ Observations in Support of Community Knowledge and Discourse

Rebecca Cober¹, Colin McCann¹, Tom Moher², Jim Slotta¹

¹Ontario Institute for Studies in Education, University of Toronto
²Department of Computer Science and Learning Science Research Institute, University of Illinois at Chicago
rebecca.cober@mail.utoronto.ca, colinmccann@gmail.com, moher@uic.edu, jslotta@oise.utoronto.ca

Abstract: We present two case studies of scientific inquiry with Embedded Phenomena, where two middle school science classes participate in whole-class investigations of phenomena that are embedded within their classroom. Students share observational data with their peers using networked handheld devices. Student-contributed data is collected, aggregated and re/presented in coherent visualizations, designed to guide students in their investigations. We examine the forms of these collective representations with a view towards understanding their efficacy in scaffolding learners and resolving their driving inquiry questions. We analyze the patterns of discourse and of use surrounding these aggregate visualizations during teacher-led, whole class discussions. For each case, we report three trends based on a visual analysis of the coded discourse, followed by a discussion of the patterns of use surrounding the display of aggregate screens. We conclude with a synthesizing discussion of the two cases.

Introduction

Inquiry-based instruction has been advocated as an approach that is well-suited for teaching and learning about science (Bransford et al., 2000). Inquiry has been defined as “an educational activity in which students individually or collectively investigate a set of phenomena—virtual or real—and draw conclusions about it. Students direct their own investigatory activity, but they may be prompted to formulate questions, plan their activity, and draw and justify conclusions about what they have learned” (Kuhn et al., 2000, pp. 496-7). In the present research, we describe two case studies of scientific inquiry concerning Embedded Phenomena (EP; Moher, 2006). EP are “ambient” media (i.e., persistent, passive, and embedded within the classroom environment) where a running simulation is mapped to— and then embedded within— the walls, floors, or ceiling of the classroom, providing a rich context of scientific inquiry for whole class investigations.

Collaborative learning is an important dimension of inquiry-oriented learning. It includes the exchange of ideas, data, or artifacts amongst peers and is often facilitated by networked scaffolding technologies (Edelson, Gordin, & Pea, 1999). A whole-class investigation approach requires that students cooperate and share data, making it likely that they will need to draw on multiple data sources, including personal- and peer-collected datasets for their analysis. Managing multiple streams of data, especially data that accumulates over several days and weeks, presents learners with a significant challenge. Computing and networking technologies have the potential to support students in such complex investigations, enabling them to efficiently collect, manage, and visualize data (Roschelle et al., 2007). Collective or aggregate representations, generated from the data gathered by individual students or small groups of students using networked devices, can take a variety of forms, including histograms such as ClassTalk (Dufresne & Gerace, 1996), annotated image maps (Tatar et al., 2003), and graphs generated from probes and sensors (e.g., Tinker, 2000).

The task of designing such representations is not simple. Although it may be relatively easy, from a technological point of view, to collect and aggregate student-contributed data, the real challenge is to present the resultant collection of data in such a way that the students’ and teacher can derive meaning from it and determine a sense of progress in their inquiry. The following research questions guide our discussion of the two cases we present:

1. Do the representations of collective input make patterns in the data visible, allowing the knowledge community to make progress towards resolving inquiry questions?
2. What interaction patterns emerge when teachers use these collective visualizations during whole-class discussions?

Theoretical Foundations

The establishment of a knowledge community in a classroom has been the aim of research projects such as Fostering Communities of Learners (Brown & Campione, 1996) and Knowledge Building (Scardamalia & Bereiter, 2006). These projects have sought to advance our understanding of the role of social interaction in
learning, emphasizing collective epistemology (Palinscar, 1998). In a learning community approach, the emphasis is on constructing knowledge through the sharing of data, ideas, and theories, within a rich social context.

Our pedagogical model, known as Knowledge Community and Inquiry, (KCI) builds on the foundation of knowledge communities as described above, with an added major emphasis on scaffolded inquiry (Slotta & Linn, 2009). In KCI, collaborative inquiry activities are carefully designed and mapped to curriculum goals, with students co-constructing a collective knowledge base, and then using this knowledge base as a resource for subsequent inquiry (Slotta & Peters, 2008). The design of KCI curriculum requires careful development of collaborative knowledge construction activities that result in a knowledge base that is indexed to the major learning goals (e.g., science standards or expectations). In addition, collaborative inquiry activities are designed to make use of that knowledge base, advance the community’s understandings, and lead to assessable outcomes.

To investigate the KCI framework, we have advanced the notion of a “smart classroom”, which is concerned with the coordination of the flow of people, roles, groups, activities, and materials, using specified pedagogical structures and curriculum content (Slotta, 2010). In this model, students and teachers work together as members of a knowledge community, using a variety of technologies, including laptops, tablets computers, interactive tabletops, and large format displays. A suite of custom software applications allows for the delivery of all materials, data collection (e.g., of student reflections, observations, tags or any other interaction), and coordination of all collaborative inquiry conditions. The smart classroom infrastructure ensures that (1) all devices connect and communicate with each other via a wireless network, (2) all inputs are appropriately attributed, sorted and available for retrieval, and (3) the developing knowledge base is accessible as a resource to individual students and the community as a whole.

The present research was conducted within an instructional environment referred to as Embedded Phenomena for Inquiry Communities (EPIC; Moher & Slotta, 2012), where KCI was applied as a pedagogical model to develop a knowledge community for elementary students to investigate Embedded Phenomena. In EPIC classrooms, students work collaboratively (i.e., in small groups) and collectively, sharing information and solving problems. Interactions, including the exchange of data and theories, are carefully designed to support the growth of collective knowledge concerning the EP under investigation, as captured in various representational forms (i.e., “aggregate representations”). In the present paper, two EPIC curricula will be described, with synthesizing discussion. In both cases, our goal was to support the construction of a community knowledge base, aggregated from the observational data collected by individuals and small groups, and to make the results of individual inquiry actions, such as observations and hypotheses, persistent and searchable. Further, we sought to dynamically generate (i.e., in “real time”) views or representations of this aggregate knowledge that would make important patterns visible to student and teachers. Using a co-design methodology (Roschelle, Penuel, & Shechtman, 2006), researchers, technologists, and teachers worked together to develop the technology-enhanced materials.

Case One: HelioRoom

Methodology

Research Setting and Design Considerations
Our first EPIC case employed the HelioRoom EP, which maps the orbital planetary system onto the four walls of the classroom. Students adopt a heliocentric perspective and observe representations of the planets (colored circles, see Figure 1) though four “windows” — monitors affixed to each wall of the classroom. As the “planets” orbit in a counter-clockwise direction, the circles appear to move off of one “window” and then reappear on the next after varying time intervals (i.e., depending on the velocity of the circle). Students track the occlusion relationships of the colored circles (i.e., when one colored circle passes in front of a different colored circle), and use data aggregated over all students’ observations to advance and support their theories concerning the identity of each planet.

In a previous nine-day implementation of the HelioRoom EP, paper cards were used to capture student observations and hypotheses, which were displayed on a classroom bulletin board. Thompson and Moher (2006) reported “the density and lack of organization of the observations on the Idea Wall made it extremely difficult for students to retrieve the evidence they needed to construct argument chains to support their theories” (p. 1001). In the present study, we applied the ideas of KCI to develop a 90-minute lesson where students worked collectively to observe all occlusion relations at all four monitors, adding their observations to a collective set using networked mobile devices (i.e., tablets) – effectively replacing the paper supports of previous HelioRoom studies – and then reflecting on the aggregate as it emerged in real time.
Participants
Two classes of grade six students and their teachers (our co-design partners) participated in the trial. These students and teachers belong to an elementary school that has a strong history of inquiry-based instruction, located in a multicultural urban centre in Canada. The present analysis is restricted to one of the two classes (n=12) in order to develop a rich case study of the discourse patterns that occurred around the re/presentations of student- and peer-collected data.

Method
Students had recently completed a curriculum unit on Earth and Space and were familiar with the order of the planets in our Solar system. The teacher began by demonstrating how to use the tablet software to contribute observational data to the knowledge base. Students worked in pairs to contribute observations. Our design supported viewing of one aggregate representation at a time (i.e., all observations for one planet). By touching a colored disc at the top of the screen, the tally (frequency count) for that color could be viewed – see Figure 1. The teacher used the interactive white board (IWB), which displayed the aggregate data, to inform whole-class discussions. Students could also view these data visualizations on their own tablets at any time during the inquiry process, and all such representations were updated in “real time” with each new observation.

Data Sources
Data sources include data logs of student contributions, a transcribed video recording of the learning activity, researchers’ field notes, audio recordings of two debrief sessions with the teachers (one following each lesson), and focus group interviews with teachers and students following the lessons.

Analysis and findings
Students contributed 425 relational observations, with an 86% accuracy rate (i.e., they correctly observed an occlusion relationship for one planet pair).

This paper analyzes the whole-class discussions that were concerned with the aggregate representations of student-contributed observational data. This analysis informs an understanding of how the knowledge community understood the collective data, as well as its role in guiding them towards their inquiry goal.

A grounded theory approach (e.g., Creswell, 2008) was used to establish a coding scheme for the statements made during the whole-class discussions: summative, interpretive, directional, negotiation, and agreement, as shown in Table 1.

Table 1: Code name, description, and example for each of the five codes

<table>
<thead>
<tr>
<th>Code Name and Description</th>
<th>Example from transcript</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summative statement regarding what the knowledge community has observed, giving words to the data</td>
<td>Teacher: “There has been one observation made that red is in front of orange.”</td>
</tr>
<tr>
<td>Interpretive statement concerning the numbers in the visualization, as well as best practices, efficacy, or accuracy of the aggregate representation</td>
<td>“Well, it depends how close the two numbers are...”</td>
</tr>
<tr>
<td>Directional statement guiding subsequent inquiry (i.e., for further observation focus) or interpretation of the aggregate (i.e., what we are looking for in the aggregate)</td>
<td>“So let’s actually look for those [areas of disagreement in the aggregate].” “Let's get some more observations of green.”</td>
</tr>
</tbody>
</table>
Patterns in Coded Discourse

Our coding of discourse that occurred relating to the aggregate representations suggests that those summary views helped to make patterns visible and allowed the knowledge community to make progress towards their goal of deciphering the identity of the obfuscated planets. The codes were plotted on a timeline (Figure 2), which includes the duration of the whole class discussions and the points at which the knowledge community reached agreement about the identity of a planet.

Patterns of Interaction

In addition to the coding of discourse, we examined when the discussions of aggregate data occurred during whole class discussions, and how they proceeded. The timeline reveals that discussions (shown as black bars in Figure 3) were interspersed throughout the learning activity at regular intervals (beginning, middle, and end). In the first two discussions, the teacher worked systematically through the aggregates by tapping the colors at the top of the screen, from left to right (see Figure 2). This provided an overview of the student-contributed
observations, allowing the knowledge community to get a sense of which planets they had already identified and those that were still unidentified. In the latter discussions, the yellow and green datasets were the focus of discussions, possibly because these were the slower moving planets, and were thus more difficult to identify because nearly every color passed in front of them. In all discussions, the teacher actively chose which of the aggregate representations would be the most productive topic for discussion, and used them to help motivate and inform students’ subsequent inquiry. In summary, earlier discussions focused on the clear emerging patterns for the fast-moving planets (presumably because these were pedagogically advantageous for the teacher), and later ones addressed the problematic circles, whose evidentiary body was sparse.

Case Two - Wallcology

Methodology

Research Setting and Design Considerations
The second EPIC case employed the WallCology EP as the setting for whole-class inquiry, targeting life sciences topics of biodiversity and population ecologies (Moher, Uphoff, Lopez-Silva, & Malcolm, 2008). Over several weeks, grade five/six students observed a digital ecosystem consisting of dynamic animations of insects and vegetation, visible through display monitors called “Wallscopes.” (see Figure 3). The ecosystem comprised four differentiated but interconnected habitats, one on each wall of the classroom, which varied in terms of environmental conditions (temperature, light and humidity). In our EPIC activity, students made observations about the morphologies and behaviors of organisms to determine their life cycle relationships. Constructing a representation of the lifecycles of any species was a challenging task; it was not always clear which organism belonged to which species (e.g., does the adult form of the “green bug” hatch from the white egg or blue egg?). It required careful observation (and maybe a bit of luck) for students to actually “see” life events unfold (e.g., laying and hatching). Additionally, since each monitor displayed a different habitat, students at one monitor see something different than students at another monitor, necessitating the sharing of observational data.

In previous implementations of WallCology (see Moher et al., 2008) students had recorded their notes and sketches in paper-based field guides, which formed the basis of whole-class discussions. The goal of the EPIC design innovation was to create a more powerful means of sharing and working with these observations, by aggregating individual or group inquiry actions, encouraging teacher and students to attend to interesting patterns in the data, including where more work is needed. Ideally, our aggregate views would reveal the most constructive patterns. We designed several applications (e.g., a graphing tool, and a modeling tool), however the present paper analyzes the role of one tool called “Life Cycle Relationships” and a corresponding aggregate representation that tallied all the pair-wise relationships observed by students using the tool.

Participants
Forty-two students from two grade five/six classes participated, with the classroom teacher guiding each class. The present analysis is restricted to one of the two groups (n=21) and to one lesson within the nine-week unit in which students contributed observations concerning life cycles.

Method
Here, we analyze how the aggregate representations of students’ observations of the lifecycle relationships of one species – the “blue bug” – helped them come to evidence-based agreements about its complete life cycle. Students had already formed two competing theories about the “blue bug’s” life cycle in a previous class, and they were interested in pursuing this line of inquiry. We hoped that such disagreements could be resolved by teacher-led discussions of the aggregate view. In this example, we hoped the teacher could help students understand which additional observations would be needed about the “blue bug’s” life cycle, and help the classroom community procure those observations and come to a consensus.

Six pairs of students and 8 single students used the tablet application to contribute life cycle observations. To contribute a life cycle observation, students selected two icons on the tablet interface — the first icon, to indicate an early life cycle stage (e.g., an egg) and the second icon, to represent the stage that immediately follows the first icon (e.g., the larva form), as shown in Figure 3. To stipulate that no life cycle stage immediately precedes the second icon (e.g., a vegetation icon), students placed the “X” icon in the first space. By tapping a life stage icon in the collective dataset view, a student could view tallies of observations concerning that organism. For each icon, there are eleven possible relational statements (e.g., “x organism precedes y organism” in a life cycle). By choosing one icon from the set of 11 icons, students and teachers view up-to-date relational tallies for the selected organism and the stage immediately preceding it, shown in Figure 3.
Figure 3. Left: Students observe WallCology ecosystem. Middle: Tablet interface used to contribute a life cycle observation. Right: Life cycle aggregate displays tallies of every observation made for the selected organism.

Data Sources
Data sources included data logs of student contributions, a video recording of the learning activity, researchers’ field notes, and focus group interviews with teachers and students following the sessions.

Analysis and Findings
Prior to this class, students had contributed 75 life cycle observations, with a 62% accuracy rate. By the end of this class period, 201 life cycle observations had been contributed, with a total accuracy of 72% — more than doubling the observations in the evidentiary database and resulting in a 16% increase in accuracy. Observations concerning the life stages of the “blue bug” were entered at a higher rate, with 37 existing in the knowledge base before the class and 113 by the end of the period. The total number of correct observations about the life stages of the “blue bug” increased threefold, and the total number of incorrect observations about it increased by 69%.

We applied the coding scheme from case one (see Table 2 for definitions of codes) to the discourse of case two, as shown in Table 2. These codes were again plotted on a timeline (shown in Figure 4), which also shows where whole class discussions occurred within the context of the learning activity and the point at which the knowledge community reached agreement about the life cycle of the “blue bug”.

Table 2: The table provides a description of the codes with an example for each, taken from the transcript.

<table>
<thead>
<tr>
<th>Code Name</th>
<th>Example from transcript</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summative statement</td>
<td>“And three people said the shrimp turned into the blue beetle.” - Student</td>
</tr>
<tr>
<td>Interpretive statement</td>
<td>“I don’t know if six people meant that the blue beetle hatches from the egg or laid the egg...” - Student</td>
</tr>
<tr>
<td>Directive statement</td>
<td>“Can we just check that that’s observed by other observers?” - Teacher</td>
</tr>
<tr>
<td>Negotiation statement</td>
<td>“So could we say for today that our current understanding of this is that the life cycle begins with the blue egg?” - Teacher</td>
</tr>
<tr>
<td>Agreement statement</td>
<td>YES!!! - Students</td>
</tr>
</tbody>
</table>

Patterns in Coded Discourse: Blue Bug discussion.
The timeline shows that the focus of the first whole class discussion was on the students’ disagreement about the “blue bugs” life cycle. The teacher said, “So what it seems like, if I’m reading the flavor of the room right, is that people have different understandings and have observed different things about the cycle of the blue bug.” The timeline indicates that five minutes before the end of the class, the issue was resolved when the knowledge community came to agreement. The coded discourse reveals three interesting patterns: (1) A directive statement was always made by the teacher (pink) when the aggregate screen was first displayed. For example, “Can someone give words to this set of data - maybe even in words like ‘most’, ‘some’, like ‘most observations’, ‘some observations’. What would you say, [student]?” (2) Students made summative (yellow) and interpretive (orange) statements following the teacher’s directive statements. Viewing the tally for the “blue bug’s” larvae hatching from its egg (24 observations), a student said, “I think almost everybody thinks that the ‘shrimp...
hatches from the egg”. Interpretative statements outnumbered summative statements, by nearly 50%. (3) The activity culminated with many negotiation statements followed by one final agreement statement: “YES”.

Patterns of Interaction
The teacher began the discussion by presenting the aggregate view of the adult stage of the life cycle. This revealed that 15 observations were made that the pupa transformed into the adult form, while only three observations were made that the larva transforming into the adult form. The teacher put forward the following hypothesis: egg \(\rightarrow\) larva \(\rightarrow\) pupa \(\rightarrow\) adult and said, “Can we just check whether that’s confirmed by other observers?” This statement provided direction for the presentation order of the aggregate screens: egg, larva, pupa, and adult. Student-contributed data was used to confirm the hypothesis, logically and systematically, resulting in resolution.

Discussion
In both cases, our designs provided the knowledge community with an overview of the “lay of the land” of their observations: a quantitative display of what the whole class had observed. Highlighting disagreement and gaps in the dataset made it possible for the community to pursue their inquiry goals, through productive discourse and by filling in obvious gaps through targeted observations.

There were also some differences that appeared between the discourse-analysis in the two cases. While the discourse codes applied equally well to both cases, there were some differences in their distribution. In the first case, many more summative statements were made, possibly because those students had never seen such a display (i.e., of binary data in table form) before. In the second case, students had already used the Life Cycle Relationships Tool in a previous class. Also, in the first case, agreement was reached several times (with negotiation and agreement statements interspersed throughout the timeline), whereas in the second case, only one issue was under discussion, and those codes appeared only at the end of the timeline.

In a follow-up interview with the teacher after the WallCology unit, he spoke of the power of aggregated data in this type of scientific inquiry. He said such representations really helped his students to consider data, as opposed to only pursuing theories (which is something his students are more comfortable with). He noted that without technology, it might be possible (although awkward) to construct such data sets, but that with networked technology it is possible to “[have] a set of data that people have contributed to, [which is] then accessible to the community in their work.” He said students felt a greater sense of accountability to each other, because they knew they were dependent on the aggregate for their information.

References


Acknowledgements
We acknowledge members of the Learning Technologies Group at the University of Illinois at Chicago and members of the Encore Lab at the Ontario Institute for Studies in Education in Toronto, who were integral to these projects. We sincerely thank the teachers and students from the Dr. Eric Jackman Institute for Child Study who participated. The material presented here is based on work supported by the U.S. National Science Foundation under grant IIS-1065275 and Canadian Social Sciences and Humanities Research Council under grant 410-2011-0474.