# Supporting Learners in Collecting and Exploring Data from Immersive Simulations in Collective Inquiry

**Michelle Lui[1], Alex Kuhn[2], Alisa Acosta[1], Chris Quintana[2], James D. Slotta[1]**

OISE / University of Toronto
252 Bloor Street West,
Toronto, Ontario M5S 1V6 Canada
{michelle.lui, alisa.acosta}@utoronto.ca,
jslotta@oise.utoronto.ca

School of Education
University of Michigan
610 East University
Ann Arbor, MI 48109 USA
{kuhnalex, quintana}@umich.edu

## ABSTRACT
Digitally augmented physical spaces (e.g., smart classrooms) offer opportunities to engage students in novel and potentially transformative learning experiences. This paper presents an immersive rainforest simulation and collective inquiry activity where students collect observational data from the environment and explore their peers' data through large visualization displays and personal mobile devices. Two iterations of the design were tested, which resulted in higher quality student explanations constructed. Images were found to be an important source of evidence for the explanations, more so than text-only evidence. We also found that patterns of collective ideas influenced student performance, and that visualizations, as ambient or plenary displays, supported both teacher and students in reviewing patterns of collected data.

### Author Keywords
Digitally augmented physical spaces; smart classroom; science inquiry; visualizations; large displays; multi-device environments; mobile computing

### ACM Classification Keywords
H.5.2 [Information Interfaces and Presentation]: User Interfaces - Interaction styles; K.3.1 [Computer Uses in Education].

## INTRODUCTION
The design and research of technology-enhanced environments can offer new opportunities for learning and instruction for K-12 science. The primary goals of these environments are to help students develop deep understanding of complicated concepts while also attaining scientific reasoning, critical thinking skills, and collaboration and communication skills [19]. In scientific inquiry, an established approach advocated by educational research, students typically investigate a phenomenon and draw conclusions about it. Instead of mastering disconnected facts, this approach places a heavy emphasis on posing questions, gathering and analyzing data, and constructing evidence-based arguments (e.g., [4, 10, 12]). However, in traditional inquiry-based learning, students typically work autonomously as individuals, pairs or, at most, in small groups. Current means of supporting inquiry instruction with technology also tend to isolate students, often confining them to work together on single machines [29].


*CHI 2014,* April 26 - May 01, 2014, Toronto, ON, Canada.

ACM 978-1-4503-2473-1/14/04…$15.00.
http://dx.doi.org/10.1145/2556288.2557162

In recent years, researchers have begun to reconsider the role of the physical learning environment and to experiment with learning activities in digitally augmented physical spaces (i.e., mixed-reality environments). These spaces offer new ways of engaging students with science concepts that have traditionally been taught or addressed through more passive forms of instruction. By coupling digital content with more familiar forms of physical engagement, students are prompted into active learning, which encourages students to think, reflect, and drive their own understanding [22]. Early efforts of digitally augmented physical learning spaces have shown positive outcomes in facilitating creativity and reflection (e.g., [1, 6, 25]). It has been suggested that coupling digital augmentation with physical spaces promotes active learning by increasing the students' awareness of the activities and information presented, provides a richer experience by bridging various perspectives between the physical and digital worlds, and engages students through a juxtaposition of familiar actions with unexpected digitally augmented results [22].

This paper explores a new form of digitally mediated "smart classroom," [30] where students are supported with personal learning devices that guide their exploration of an immersive simulation (Figure 1). The environment is responsive to student observations (recorded on mobile devices), with real-time emergent visualizations that serve to capture and aggregate student observations for knowledge building and discourse.

Our research is concerned with the design of inquiry activities that complement and define such immersive environments, where students are engaged as a whole class, jointly negotiating problems and working towards a common goal. In this form of inquiry-based learning, called

**Figure 1. Students engaging in EvoRoom inquiry activity.**

"collective inquiry," the entire classroom of students is treated as the focus of interaction design. Collaborative inquiry—a related form of inquiry-based learning—in contrast, tends to focus on interactions between pairs or small groups of students (e.g., [29, 26]). In collective inquiry, students are encouraged to think deeply about materials and develop their own understandings, but with an emphasis on collective knowledge or progress over individual contributions. Students are responsible for generating and building upon each other's ideas, ultimately developing and depending upon their own community knowledge base [21, 28]. Our motivations for using digitally augmented physical spaces for advancing a collective epistemology is to leverage more natural forms of social interactions afforded by such spaces, to enhance collaboration and authenticity of the learning experience.

While considerable research has addressed the design of digitally augmented physical spaces, few projects have investigated rich, immersive experiences for inquiry-based learning, and none have investigated collective inquiry in digitally augmented physical spaces. This paper explores how the physical space of the room can serve to mediate students working as a knowledge community to collect data, support the visualization of the community's aggregate data, and facilitate inquiry activities, including discussions of data. Specifically, this paper considers the following research questions over two design iterations:

RQ1: How do we support learners in collecting observational data from immersive simulations in a way that promotes scientific reasoning?

RQ2: How do real-time visualizations on large shared displays support learners and teachers in assessing and discussing the aggregated observational data?

## BACKGROUND

### Digitally Augmented Physical Spaces for Learning

Much of the prior work on using embedded, ubiquitous computational media to support learning is based on the pioneering work of Wilensky and Resnick, who argued for the use of pervasive, non-desktop technologies that engage learners with the physical environment [35]. The concept of participatory simulations [3], where students themselves become the elements of the simulation (e.g., with the use of physical manipulatives), demonstrates this idea using role-playing activities. Colella [3] for example, gave students wearable transmitting badges, transforming them into potential virus carriers with the mission of greeting as many classmates as they could without getting "sick."

The Hunting of the Snark [25] combines the use of mobile technology with the physicality of the classroom environment. Young children (ages 6-10) investigate a virtual creature by exploring a variety of physical spaces the creature might be located (e.g., cave where it sleeps). Location tracking devices, pressure pads, and other sensor technology is used to engage pairs of children in interacting with the creature (which responds with sounds) and exploring its characteristics. The children in the study generated theories about the Snark (i.e., describing its emotional state), and were found to engage in various forms of collaboration to creatively experiment and explore ways of interacting with the Snark [23].

The Embedded Phenomena (EP) framework leverages the physicality of the classroom, carefully "mapping" a persistent scientific simulation onto the walls or floor of the room in the form of location-specific computer "portals" [16]. Students monitor and manipulate the simulation and gather evidence to solve a problem or answer a question. The use of EP was associated with greater learning gains compared to students who experienced an analogous "non-embedded" curriculum (i.e., the simulation activity was replaced with an exercise on desktop computers) in a quasi-experimental, within-teacher study [17].

Another example of digitally augmented learning spaces is that of SMALLab, where high school students studied geologic evolution by collaboratively constructing and monitoring the earth's crust, investigating uplift and erosion over time [2]. Using various input devices (e.g., glowballs, Wii remotes, wireless game pads) and a projected interface on the ground, groups of students were responsible for building, maintaining or evaluating a cycle of the geologic clock. The intervention resulted in significant achievement gains, demonstrating the promise for further research regarding face-to-face interactions in a computationally-augmented physical space, and distributed roles through a generative process that unfolds over time [2].

The examples above illustrate the potential of such innovative technologies for learning. However, there remains a need to understand how digitally augmented physical spaces can contribute to collective approaches in classroom learning. In particular, we are concerned with how to capture knowledge, as it emerges in real-time, including aggregate or summative aspects (i.e., "wisdom of the crowd") as well as discrete knowledge elements—all the while carefully balancing face-to-face interactions

**Figure 2. EvoRoom configuration. Consisting of six projected displays (three per side) and two interactive whiteboards (middle).**

within inquiry activities. One important challenge involves designing interfaces that capture and represent student contributions, and making them accessible to peers.

### Capturing and Visualizing Student-Collected Data

Researchers have used a variety of technology solutions to support learners in collecting and sharing their observations with peers. Several researchers have identified the need for visualizations to give users an overview of data while also supporting requests for details on demand [27]. In the Participate project, data was collected by and shared amongst students (ages 13-15). Using Google Earth and graphing software in conjunction with GPS enabled data collection equipment and mobile phones, students explored visualizations of large numerical datasets [36]. This research found that student generated data was an engaging source of discussion and inquiry in later classroom activities. Fraser's SENSE project, in which students (ages 10-14) collect environmental data in the field and share this information with their peers on returning to the classroom, found the process of viewing data from another class helped students reflect and understand new perspectives [7].

The Zydeco project [11] used a mobile app to support students in collecting data from the field and found that students were able to collect data from informal environments and have the students successfully evaluate and utilize their personal and peer-collected data to construct science explanations. Students were able to review and filter their data set by the different ways the data was annotated (title, tags, questions the data was collected under, and media types present in the data).

Interestingly, none of the projects described here employed large displays for the visualization of student data. Rather, they relied on teacher-led class discussion based on the display of aggregated data presented on individual devices.

The use of physically shared displays for Computer-Supported Cooperative Work (CSCW) has been advanced in the development of single display groupware (SDG) [31, 33]. SDG emphasizes face-to-face interactions, enabling multiple co-located users to collaborate via a single shared display using multiple (and simultaneous) input devices. Early applications in educational settings, such as KidPad [5] and Give and Take [9] investigated how small groups of children could work together creating stories and solving puzzles respectively. More recent efforts focus on supporting larger groups of children (e.g., [18] who engaged up to 32 students), or groups of children playing collaborative learning video games [8].

With the proliferation of personal mobile devices, a natural extension of the design solution to support co-located collaboration are multi-device environments (MDEs), where multiple personal and shared devices are connected with specialized underlying software architecture [34]. Large shared displays from such environments have been shown to improve collaboration by promoting group awareness and communication processes [34], suggesting possible applications for collective inquiry. This paper investigates the use of large shared displays to advance collective knowledge work.

## EVOROOM

In order to understand how digitally augmented physical spaces can contribute to collective approaches in classroom learning, our research group at the University of Toronto developed a "smart classroom" research environment at a partner secondary school. Previous reports described the smart classroom technology framework [28], and learning outcomes in its application in mathematics and physics activities [14, 32].

EvoRoom represents the first instantiation of immersive simulations designed for the smart classroom environment. Students are situated in the rainforest of Borneo through two sets of large projected displays (achieved by stitching together three displays) that students examine during collective inquiry activities designed in parallel with the immersive simulations (Figure 2). Two interactive whiteboards are located at the front of the room. The simulation files are networked and controlled with a custom application that allows the teacher to manage the time spent in each portion of the activity, controlling the pedagogical flow within the room.

EvoRoom was co-designed [20] with a teacher to fit seamlessly within a broader secondary school biology curriculum, in topics of evolution and biodiversity. Running for 10 to 12-weeks, the integrated curriculum includes in-class activities, homework, a field trip to the zoo, and two collective inquiry activities with immersive simulations.

To date, a pilot study and two design iterations of EvoRoom have been developed and tested. During the collective inquiry activities, students take on the role of "field researchers," and work individually, in small groups, and as a whole-class to complete tasks delivered to them on their personal tablet computers. The tablets help to place students in small groups, scaffold their activities, collect observations, and give real-time updates and resources. Student observations and reflections are aggregated and displayed on the interactive whiteboards in real-time. In the first design iteration, a custom EvoRoom tablet application was developed to support the collection of student observations. In the second iteration, Zydeco [24] was used to collect and review student observations, as it offered more data collection and sharing capabilities.

The first collective inquiry activity focused on evolutionary concepts. Students gathered evidence of evolution by observing changes in life forms as the simulation advanced across two hundred million years. Findings from the pilot study and the first design iteration of this activity were previously reported. Early designs of EvoRoom were found to engage students, and curricular scaffolding in the first full implementation was tied to increased variation and complexity in student ideas about evolutionary topics [13].

The present paper focuses on the second collective inquiry activity, on the topic of biodiversity. Prior to the activity, students wrote predictions about how changing an environmental factor (e.g., tsunami, earthquake, low rainfall) for a single season could change the biodiversity in Borneo over a five-year time span. In the immersive environment, students were presented with four different versions of the Borneo rainforest ecosystem. The activity challenged them to explore the differences between these four rainforests and to locate the rainforest that resulted from the variable they explored in their predictions, using the following digitally mediated interactions:

1. Engaging with immersive simulation: Students go up to the walls and find evidence about changes to the ecosystem that might have taken place;
2. Engaging with mobile devices: Students make observations and record their findings using tablet computers. Students also work in small groups, with interactions scaffolded by their tablets;
3. Engaging with real-time emergent visualizations: Student and teacher interactions form around large emergent displays of student observations.

We have previously described early designs of EvoRoom [13], including how the immersive simulation and collective inquiry activities impact student ideas about evolutionary concepts. Here, we report on a second iteration of the collective inquiry activity on biodiversity. This work builds on previous reports on student learning by: 1) examining how student interactions with personal devices and large emergent displays are associated with learning outcomes, and 2) their implications on how specific design features might impact the learning experience.

## METHOD

Our core design team, consisting of two researchers and a high school biology teacher collaborated using a co-design methodology to create the simulation, inquiry activity and interactive materials [20]. Our co-design partnership began a year prior to the first enactment. Meeting approximately once per week during the academic school year, we considered important design elements and outlined our overall strategy. In the months leading up to the enactments, design meetings widened to include two technology developers, and (in the second trial) three additional researchers.

Participants included students (ages 14-16) from four class sections of Grade 11 Biology taught by our co-design teacher. The first iteration included 45 students. The second trial occurred in the following academic year, included 54 students. All of the participants (99 students) entered the study with no prior experience of the immersive simulation.

In both trials, video and audio recordings served to capture patterns of interactions within the room. All student artifacts were collected for analysis. Two members of the research team coded students' observation notes, with discrete "observations" serving as a unit of analysis. The observations were judged for whether they could potentially be useful to the investigation or not [11]. An example of a useful observation is: *"The jambu tree, rafflesia, and ginger are all missing."* An example of a scientifically unproductive observation is: *"Scorched earth indicative of magic."* The quality of students' final explanations were analyzed using the Claim-Evidence-Reasoning (CER) grading rubric [15], which was modified slightly to fit the context (Table 1).

## TRIAL 1 DESCRIPTION

Prior to visiting EvoRoom, students worked in groups of two or three to complete their predictions assignment on

| Claim | Evidence | Reasoning |
|---|---|---|
| 0: Nothing/not a claim | 0: No evidence | 0: Reasoning absent, non-sensible, or fails to connect evidence to claim. |
| 1: Claim is incomplete or not related to the question | 1: Does not provide relevant evidence | 1: Provides reasoning that links the claim and the evidence by repeating the evidence |
| 2: Makes a complete claim related to the question | 2: One relevant evidence | 2: Explains why evidence supports the claim, but does not address all aspects of |

**Table 1. Modified Claim-Evidence-Reasoning rubric.**

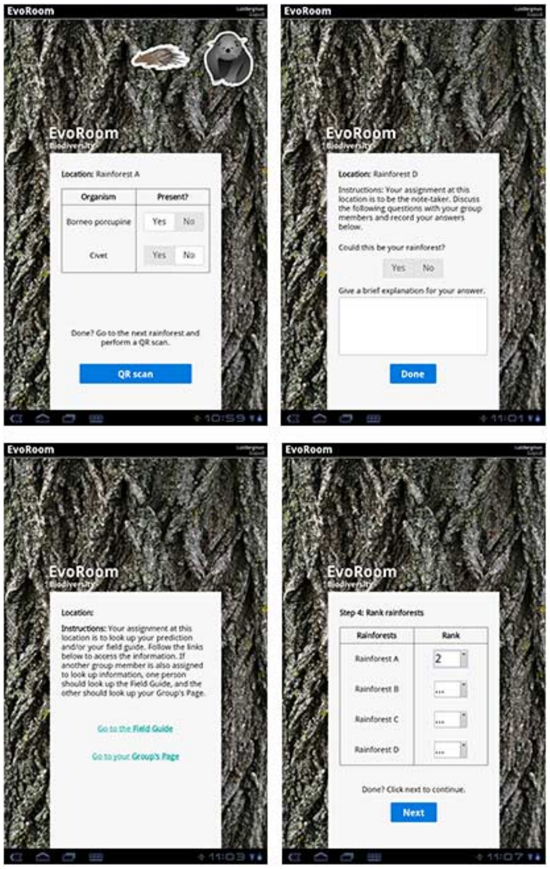


**Figure 3. Custom tablet interface from trial 1.**

their class website. Each group also commented on at least two predictions made by other groups.

Four sessions (each with different students) were run, with 10 to 12 students split into four groups. Each student was provided a tablet computer with a custom designed application that supported navigation through the activity while scaffolding students to work together and to collect data throughout the session (Figure 3).

The activity was designed for a 75-minute class period, including several data collection steps where students examined the organisms in each of the four rainforest stations, eliminating rainforest stations that they felt did not result from their factor, and discussing their findings with other groups. Using the tablets, students contributed written explanations to simple questions (e.g., *"Could this be your rainforest?"*). All of the groups' decisions and notes were displayed on the interactive whiteboard at the front of the room. Once data collection was complete, students ranked the four rainforest stations according to which was most likely caused by their assigned factor. Students explained their group's choice by answering follow-up questions (e.g., *"What is your strategy in ranking the rainforests? Describe your process"*). The aggregated results were shown on the interactive whiteboards, and were also used by the teacher in their classroom the following day, for a final discussion.

## TRIAL 1 RESULTS

### Scaffolded data collection

To evaluate how learners collected observational data from immersive simulations, we examined the characteristics of their initial hypotheses, which were written about each rainforest station as a group. A total of 146 discrete observations were coded from 65 notes, with each group contributing an average of 9.1 observations (*SD*=2.2). Inter-rater reliability (IRR) was performed on 20% of data, where 92.3% agreement (Kappa=0.90, $p<0.001$) was achieved. 145 (99%) of the observations were coded as being potentially useful indicating that students tended to be engaged in the task.

Overall 6 out of the 16 total groups correctly identified the rainforest stations that were associated with their assigned environmental factor. In one session, all four groups correctly identified their climatic scenario. Generally, students who made the correct match used a specific strategy, such as process of elimination or identifying a set of characteristics that matched their predictions. The explanations for the groups' final choices were evaluated, with average scores of 6.0 of 8 (*SD*=1.4). IRR was performed on 20% of data, where 100% agreement was achieved. A one-way ANOVA was conducted to evaluate the relationship between the groups' CER scores and their identification accuracy, which found that students who correctly identified rainforest stations received significantly higher CER scores than groups who incorrectly identified the rainforest stations, $F(1, 14)=7.30$, $p<0.02$.

### Interactions supported by emergent visualizations

85% of students referred to the visualizations at the interactive whiteboards (Figure 4. Students also interacted with the whiteboard to obtain further details of the data being displayed, which was done by one student in three of four sessions and four students in one session. The teacher interacted with the interactive whiteboard the most, with an average of 8.3 interactions per session (*SD*=2.9). She reviewed and organized student work, using the displays to monitor student progress.


**Figure 4. Emergent visualizations from trial 1.**

Additionally, ad hoc discussions were observed between the teacher and students near the interactive whiteboards. When the teacher was located at the boards, students would approach, either curious about the display, or wanting to consult with teacher. They would engage in discussion, with the teacher often gesturing towards the boards and using them as an aid to answer the students' questions. When students independently walked up to the displays without the teacher's presence, their group members often joined them and engaged in their own discussion. In most such cases, the teacher then joined the group discussion.

## DESIGN ITERATION

### Scaffolded data collection

For the second iteration, an important goal was to encourage more thoughtful interaction with the immersive simulation. We felt that in the first trial, while we were satisfied with the number of observations contributed and the general quality of observations, we had hoped for more accurate identification of the rainforest stations. One idea we wished to explore further was taking photos of the immersive visuals, as a mode of data collection, which could potentially support students in making more intentional comparisons between the rainforest stations.

*Design change 1:* Allow students to collect evidence in the media of their choice (e.g., photo, video, audio note, text note) and include a "control" version of Borneo (with no environmental impact), to promote more thoughtful comparisons of the immersive simulation.

From the first trial, when we examined students' recorded explanations regarding how they chose the rainforest (i.e., strategy, supporting evidence used, additional information), we noticed that the nature of the evidence ranged from lists (e.g., *"No water. Trees losing leaves. Low flora and fauna."*), to detailed explanations for all stations observed (e.g., *"Rainforest A has too much water. Rainforest B seems hot and hazy, hotter than rainforest D, and if there was an extreme increase in temperature, so B is higher than in our preference. Rainforest C has abundant life and has a lot of biodiversity, so if there was a slight increase in temperature, the plants could possibly flourish more"*). In some of the responses, we were unable to determine the evidence underlying their explanation (e.g., *"Looking at plant life and animal life"*). This prompted us to revise the data collection design to require students to explicitly link one or more of their (or their peer's) observations within their final explanations.

*Design change 2:* Have students link observational data in their final explanation as evidence, and structure observations using annotations (e.g., title, tags, questions) that denote environmental factors (e.g., high rainfall) and station identification (e.g., station A) for students to sort the data.

### Interactions supported by emergent visualizations

The emergent visualizations presented in the first trial were in the form of ambient displays, which students could review on their own. However, we noted that not all students reviewed the front board and thus did not consider the collective data their peers had collected.

The feedback from the teacher was that the interactive whiteboard was a useful resource for spontaneous and planned discussions. Hence, for the second trial we set aside time for the students to formally review the visualization and have a group discussion, facilitated by the teacher, to help everyone benefit from reviewing the collective work.

*Design change 3:* Inclusion of a teacher-facilitated data review step with emergent visualizations to support students in understanding their collective knowledge base.

### Incorporating Zydeco with EvoRoom

Zydeco offered several features we hoped to incorporate in our tablet application: 1) the ability for students to collect observations in a variety of media forms (e.g., images, audio, written notes); 2) the scaffolding of data collection and analysis of observational data through the use of annotations; and 3) the ability to link existing observations to student explanations. We opted to incorporate Zydeco into the second design iteration rather than make significant changes to our existing custom application.

An extension of Zydeco to visualize student-collected data on the interactive whiteboards was designed and developed for the second trial. A trade-off from this was that separate tablets were needed to drive the interactivity of the boards (limiting access to the those handling the tablets).

## TRIAL 2 DESCRIPTION

As in the first iteration, students in trial 2 were assigned an environmental factor and wrote predictions about its effects on biodiversity in Borneo. Students visited EvoRoom in cohorts of 13 to 14 students and were again split into four different groups of three or four members. Within the 75-minute class period, students used a handheld computer to collect evidence concerning which rainforest station showed the effects of their group's assigned environmental factor. The teacher facilitated a data review session that revealed students' initial ideas, which was followed by students constructing explanations about their final decision. Each group displayed their work and presented their findings. At the end of the session, the teacher revealed the correct answers (i.e., which station resulted from which environmental factor) and held a deeper discussion about environmental impacts and biodiversity.

## TRIAL 2 RESULTS

### Scaffolded data collection

193 data objects were collected, with an average of 12.1 pieces of data collected by each group (*SD*=5.6). The majority (89.1%) contained an image, with less than half of all data collected containing text notes (42.5%). Of the data collected, those observations with accompanying descriptive text (in either text note or title) or audio (either in the audio note or video entry) were judged in a similar

fashion as in trial 1, where each discrete unit of observation was counted. 190 observations were coded, with each group contributing an average of 11.9 observations (*SD*=8.3). It was interesting that three groups in the same session made fewer than five observations, tending to only take photographs of areas and verbally discuss their observations but did not record it. 97.9% of the observations were categorized as potentially useful. IRR was performed on 20% of data, where 87.2% agreement was achieved (Kappa=0.83, *p*<0.001).

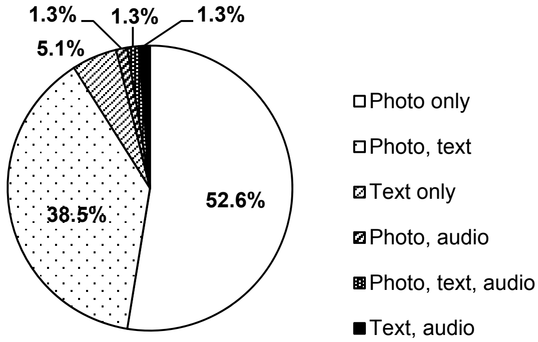


**Figure 5. Composition of data objects used as evidence.**

Similar to the results in the first trial, 6 out of the 16 total groups correctly identified the rainforest stations and in one of the sessions all four groups correctly identified their environmental factor. However, students' final explanations scores were higher in trial 2, averaging 7.6 of 8 (*SD*=0.6), with no significant relationship between identification accuracy and CER scores (as in trial 1). IRR was performed on 20% of data, where substantial agreement was achieved (Kappa=0.71, *p*<0.01). It should be noted though that in the session for which all four groups correctly identified their factors, all groups achieved perfect scores for their final explanations.

Of all the data objects that were collected, 40.4% were later used as evidence to support the explanations. Most of the of the evidence used to support claims included a photo (93.6%), compared to less than half which had a text note attached to the data object included as evidence (46.2%; Figure 5). One group mentioned that they only took photos because they had previously conducted a lot of research on the different environments and were looking for characteristics in the photos they knew would be in each environment. Another group fixated on certain traits they knew their scenario should have, and went looking for any observational evidence that exhibited those traits. For example, high rainfall would result in an environment with more water and greater biodiversity, and so they took photos of any elements of the displays with plentiful water.

### Interactions supported by emergent visualizations

Only the teacher's use of the emergent visualization was analyzed in the second trial, since the teacher used a separate tablet to drive the interactivity of the boards. In each of the four sessions, the teacher used the filtering options of the emergent visualization to ensure that students were collecting evidence for all stations and tagging evidence properly (Figure 6). During the data discussion, the teacher used the visualization's filtering functionality to review the groups' work. For example, toggling the visibility of "earthquake" observations and hiding all others revealed only those associated with earthquakes. In the visualization, the observations were organized according to the rainforest stations the observations were about. In this way, the teacher was able to notice preliminary trends (e.g., *"station B and C could be low rainfall"*). The teacher also expanded the data objects to reveal detailed information (e.g., text note, annotations), reading out examples or examining larger images. After the data discussion, students accessed the collective set of data objects from their own tablets (Figure 7).

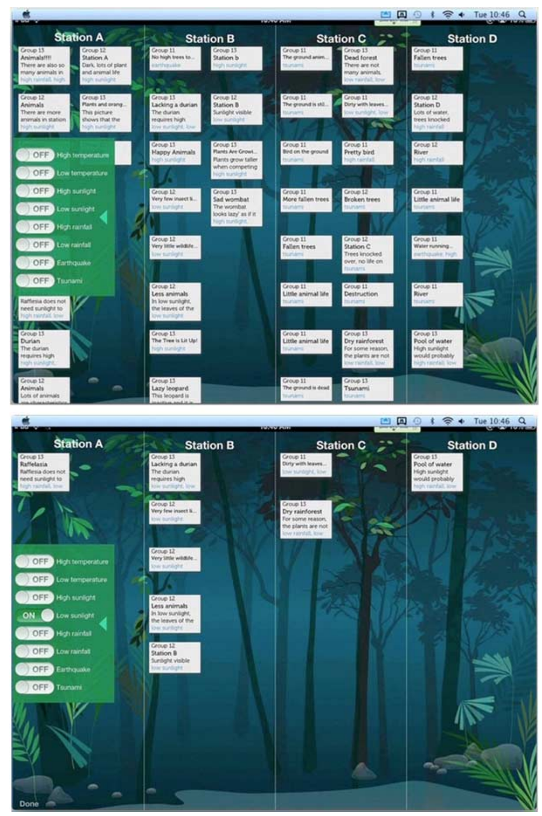


**Figure 6. Emergent visualization used in trial 2.**
Unfiltered (top) and filtered by "low sunlight" (bottom).

To understand how students used their session's collective information, we examined the students' final choice with the pattern of collected observational data tagged with their particular climate scenario. 13 of 16 groups (81.3%) chose a station associated with the highest number of data objects tagged with their environmental factor. In the students' final explanations, two of the three remaining groups indicated that they considered the collective data, but had reasons to believe that the more popular choice was incorrect.

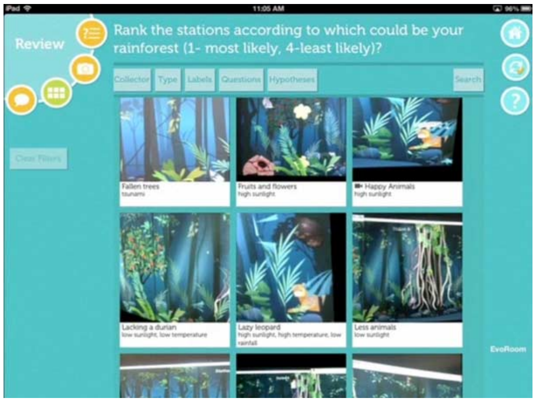


**Figure 7. Tablet display in trial 2 for reviewing peer-collected observational data.**

However, in these three cases, their final choices as well as the stations associated with the highest number of data objects tagged with their scenario were both incorrect. The highest number of data objects was annotated with the correct environmental factor and rainforest station in only 8 of 16 instances (50% collective accuracy).

## DISCUSSION

### Scaffolded Data Collection

The first design change allowed students to collect evidence in the form of photographs, video or notes. The decision was made to promote more thoughtful interactions with the immersive simulation, however the students in trials 1 and 2 achieved the same level of accuracy. There were also comparable numbers of text-based observations, recorded per group in the two trials, as well as comparable numbers of written observations that were deemed useful. However, there were a number other reasons that led us to believe that this was a positive change.

In trial 2, the majority of observations collected contained photographs. While the number of text-based observations in the two trials was comparable, there was an additional set of photographic observations in trial 2, which meant that a great deal more observations were made in the second trial. Further, a majority of trial 2's explanations (94%) relied on data objects that contain images as evidence (a form of observational data not used in trial 1), which points to its effectiveness in the immersive environment. This was somewhat surprising since we anticipated some dependence on written observations to serve as evidence, indicating that while written observations were useful to some, other students found images informative. This may be because the annotations accompanying the images supplied students with sufficient information about the observed. It could also be that the highly visual immersive simulation environment lent itself to image-based observations. By enabling access of information through different senses, we hoped to increase the awareness of students being in the immersive space, which would have provided a richer basis for reflection [22].

The CER scores of students' final explanations was another outcome that distinguished trial 1 from 2. Despite the same level of accuracy attained, the CER scores for explanations in trial 1 were lower than those in trial 2. It was also interesting that the CER scores in trial 1 corresponded with the accuracy of the students' final choice, but the same was not observed in the second trial (due to the higher average CER scores). This indicated that more students were able to produce well-reasoned claims with appropriate supportive evidence. Incorrectly identified rainforest stations were sometimes accompanied by strong explanations with appropriate evidence. This was the case if students' incoming ideas about the impact of their environment factor differed from what was presented as the correct answer. As such, the students' inability to identify rainforest scenarios only served to highlight false assumptions and possible rigid thinking, unrelated to their observational data collection. Thus, the pedagogical goal of this activity was not focused on the correct identification of rainforest stations, but rather on scientific reasoning from a collective inquiry process.

The second design change allowed students to structure observations using annotations and explicitly link data in their final explanation as evidence. For the most part, students successfully annotated the observational data. In cases where tags were not properly assigned, the teacher noticed the problematic observations in the emergent visualization and asked students to revise their notes.

In trial 1, although observational data was collected, the data were only accessible from the front boards, which were not as easy to examine. Nor did everyone review them. This led us to believe that the observations themselves were largely ignored when students constructed their explanations, although the experience of collecting the observations likely supported them in making their final choices. In trial 2, we had students spending more time reviewing the evidence they had collected through their tablets, filtering through large data sets using annotations (e.g., by searching "low sunlight") and structuring their explanations such that they had to attach existing pieces of observations to their claims. This way, we were able to see if there was any relationship between the kinds of observations that were used as evidence and their claims. As noted above, we saw an improvement in CER scores in trial 2 over trial 1 when observational data served as evidence to students' final explanations. This result may be attributed partially to both design changes—from more intentional data collection and thoughtful interaction with the immersive environment, as well as explicitly linking student observations to support their explanations.

### Interactions Supported by Emergent Visualizations

In the first trial, emergent visualizations designed to show larger patterns of early ideas fostered ad hoc discussions when used as ambient displays. The third design change in trial 2 was a teacher-facilitated data review step with the visualizations to help more students consider the collective knowledge base. The data review discussion revealed patterns of early ideas that students had, however many students had false assumptions, which permeated through to the data they collected. Showing these early patterns (that had an average accuracy rate of 50%) potentially affected students' final choice and overall accuracy.

It was interesting that in both trials there was one session where all the groups selected the correct rainforest station for their assigned factor. In the other sessions, the accuracy ranged between 0% and 50%. In the session with the highest accuracy rate in trial 2, the collective pattern shown during the data review step was essentially correct (i.e., the highest number of data objects was annotated with the correct environmental factor for each of the four rainforest stations), which likely contributed to the high accuracy rate. While showing a correct collective pattern translated to a higher accuracy rate, showing incorrect collective patterns was associated with lower accuracy. It is true that technology can help the teacher and students reveal their assumptions and have this information corrected, but it should be presented in a manner that encourages students and teacher to advance the discussion. For example, collective knowledge could be positioned to students as a body of work to be revised and built upon, rather than a "completed" state of knowledge. Future designs could benefit from providing more support and guidance in encouraging students to think about what the collective set of observational data mean for their own question in the investigation, and how this may be built upon and be used in their explanation. Large shared displays, such as the emergent visualizations, have been used as status displays to promote group awareness, while those that replicate content from personal devices have been shown to improve efficiency of collaboration by facilitating conversational grounding [34]. Our work contributes to this body of research on MDEs by advancing the design of large shared displays that serve to display both status information, as well aggregated content from numerous participants' devices (i.e., emergent visualizations).

We found that as ambient displays, the visualizations supported teachers in monitoring student progress both in terms of tasks completed, as well as depth of thinking. For students, they grounded conversations, and the formation of physical foci for ad hoc discussions with group members and with the teacher. They served as real-time representations of the collective knowledge, demonstrating emergent patterns, and served as presentation displays for plenary discussions.

### Limitations

One limitation of our work may be the over-reliance on student handheld computers, which could actually serve to disrupt naturally occurring collaborations (i.e., as students are distracted by the specific mediations of the handheld environment). Another is related to the cost of the innovation. We are cognizant that our project consists of a unique system not typically found in classrooms today, and knowing the potential costs, we have tried to set it up with hardware that schools may already have (e.g., projectors, interactive whiteboards)—with the hope that they will become more available for the uses outlined in the current paper.

## CONCLUSION

This work explored how student collected data from immersive simulations and how large visualization displays supported teachers and students in reviewing the students' collective work. It presents one of the first studies on collective inquiry through the use of immersive simulations to engage students in a traditionally challenging scientific domain through several levels of interaction (i.e., individual, small group, and whole class). It also serves as an example of a digitally augmented physical space and MDE that incorporates the use of personal devices, mixing one-to-one screen experiences with collective immersion via several large-screen displays of dynamic content and shared displays of emergent visualizations of collective knowledge work. With respect to our specific research questions, we found that:

1. Including photographs of the immersive simulations was important to students' data collection, when large collections of observational data are structured with tags or other annotations (RQ1).
2. Emergent visualizations afforded naturalistic social interactions when placed as ambient displays, but require the design of explicit interactions, to ensure active engagement of all participants. (RQ2).
3. The patterns of collective ideas presented influenced student performance, either positively or negatively, depending on the accuracy of the pattern shown. This suggests the need for careful design, including real-time monitoring and input from the teacher (RQ2).

This work revealed a number of opportunities for refinement of the EvoRoom activities and technology supports, particularly concerning how to design collective knowledge building experiences in digitally augmented physical spaces. Future research will explore additional technologies and form factors for supporting student-to-student, student-to-teacher, and student-to-whole-class interactions, and design opportunities for such interactions to influence student inquiry.


## ACKNOWLEDGMENTS

We are grateful to our partner teacher María Niño-Soto for her extensive contributions. We thank members of the EvoRoom and Zydeco team and student volunteers for time given to the project. We also acknowledge Lana Yarosh, who gave helpful guidance to our preparation of the final version of this paper. The work reported here was funded partially by the Social Sciences and Humanities Research Council of Canada Grant 410-2008-2125 and the National Science Foundation Grant DRL-1020027. Any opinions and findings expressed herein are those of the authors and do not necessarily reflect those of the funding agencies.


## REFERENCES


1. Bell, M., Chalmers, M., Barkhuus, L., et al. Interweaving mobile games with everyday life. *Proc. CHI 2006*, ACM Press (2006), 417–426.
2. Birchfield, D. and Megowan-Romanowicz, C. Earth science

learning in SMALLab: A design experiment for mixed reality. *J. of CSCL 4*, (2009), 403–421.
3. Colella, V. Participatory simulations: Building collaborative understanding through immersive dynamic modeling. *J. of the Learning Sciences 9*, 4 (2000), 471–500.
4. de Jong, T. and van Joolingen, W.R. Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research 68*, 2 (1998), 179–201.
5. Druin, A., Stewart, J., Proft, D., et al. KidPad: A design collaboration between children, technologists, and educators. *Proc. CHI 1997*, ACM Press (1997), 463-470.
6. Facer, K., Joiner, R., Stanton, D., et al. Savannah: Mobile gaming and learning? *J. of Computer Assisted Learning 20*, (2004), 399–409.
7. Fraser, D.S., Smith, H., Tallyn, E., et al. The SENSE project: A context-inclusive approach to studying environmental science within and across schools. *Proc. CSCL 2005,* ISLS (2005), 155–159.
8. Infante, C., Weitz, J., Reyes, T., et al. Co-located collaborative learning video game with single display groupware. *Interactive Learning Environments 18*, 2 (2010), 177–195.
9. Inkpen, K., Booth, K.S., Gribble, S.D., and Klawe, M. Give and take. *Proc. CHI 1995*, ACM Press (1995), 258–259.
10. Krajcik, J.S. and Blumenfeld, P.C. Project-based learning. In R.K. Sawyer, ed., *The Cambridge Handbook of the Learning Sciences*. Cambridge University Press, New York, 2006, 317–334.
11. Kuhn, A., McNally, B., Schmoll, S., et al. How students find, evaluate and utilize peer-collected annotated multimedia data in science inquiry with zydeco. *Proc. CHI 2012*, ACM Press (2012), 3061-3070.
12. Kuhn, D., Black, J., Keselman, A., and Kaplan, D. The development of cognitive skills to support inquiry learning. *Cognition and Instruction 18*, 4 (2000), 495–523.
13. Lui, M. and Slotta, J.D. Immersive simulations for smart classrooms: exploring evolutionary concepts in secondary science. *Technology, Pedagogy and Education*, (2013), 1–24.
14. Lui, M., Tissenbaum, M., and Slotta, J.D. Scripting collaborative learning in smart classrooms: Towards building knowledge communities. *Proc. CSCL 1*, ISLS (2011), 430–437.
15. McNeill, K.L. and Krajcik, J.S. Supporting grade 5-8 students in constructing explanations in science: The Claim, Evidence, and Reasoning framework for talk and writing. *Pearson Allyn & Bacon*, (2011).
16. Moher, T. Embedded Phenomena: Supporting science learning with classroom-sized distributed simulations. *Proc. CHI 2006*, ACM Press (2006), 691–700.
17. Moher, T., Wiley, J., Jaeger, A., et al. Spatial and temporal embedding for science inquiry: An empirical study of student learning. *Proc. ICLS 2010*, ISLS (2010), 826–833.
18. Moraveji, N., Inkpen, K., Cutrell, E., and Balakrishnan, R. A mischief of mice: Examining children's performance in single display groupware systems with 1 to 32 mice. *Proc. CHI 2009*, ACM Press (2009), 2157–2166.
19. National Science Foundation. *Preparing the Next Generation of STEM Innovators*. 2010.
20. Penuel, W.R., Roschelle, J., and Shechtman, N. Designing formative assessment software with teachers: An analysis of the co-design process. *Research and Practice in Technology Enhanced Learning 2*, 1 (2007), 51–74.
21. Peters, V.L. and Slotta, J.D. Scaffolding knowledge communities in the classroom: New opportunities in the Web 2.0 era. In M.J. Jacobson and P. Reimann, eds., *Designs for Learning Environments of the Future*. Springer, Secaucus, NJ, 2010, 205-230.
22. Price, S. and Rogers, Y. Let's get physical: The learning benefits of interacting in digitally augmented physical spaces. *Computers & Education 43*, (2004), 137–151.
23. Price, S., Rogers, Y., Scaife, M., et al. Using “tangibles” to promote novel forms of playful learning. *Interacting with Computers 15*, 2 (2003), 169–185.
24. Quintana, C. Pervasive science: Using mobile devices and the cloud to support science education. *interactions 19*, 2012, 76–80.
25. Rogers, Y. and Muller, H. A framework for designing sensor-based interactions to promote exploration and reflection in play. *International Journal of Human-Computer Studies 64*, 1 (2006), 1–14.
26. Rogers, Y. and Price, S. The role of mobile devices in facilitating collaborative inquiry in situ. *Research and Practice in Technology Enhanced Learning 3*, 3 (2008), 209–229.
27. Shneiderman, B. The eyes have it: A task by data type taxonomy for information visualizations. IEEE Comput. Soc. Press (1996), 336–343.
28. Slotta, J.D. Evolving the classrooms of the future: The interplay of pedagogy, technology and community. In K. Makitalo-Siegl, F. Kaplan, J. Zottmann and F. Fischer, eds., *Classroom of the Future: Orchestrating collaborative spaces*. Sense, Rotterdam, 2010, 215–242.
29. Slotta, J.D. and Linn, M.C. *WISE science: Web-based inquiry in the classroom*. Teachers College Press, 2009.
30. Slotta, J.D., Tissenbaum, M., and Lui, M. Orchestrating of complex inquiry: Three roles for learning analytics in a smart classroom infrastructure. *Proc. LAK 2013*, ACM Press (2013), 270–274.
31. Stewart, J., Bederson, B.B., and Druin, A. Single display groupware. *Proc. CHI 1999*, ACM Press (1999).
32. Tissenbaum, M., Lui, M., and Slotta, J.D. Co-designing collaborative smart classroom curriculum for secondary school science. *J. of Universal Computer Science 18*, 3 (2012), 327–352.
33. Tse, E. and Greenberg, S. Rapidly prototyping Single Display Groupware through the SDGToolkit. *Proc. AUIC 2004*, ACS (2004), 101-110.
34. Wallace, J.R., Scott, S.D., Lai, E., and Jajalla, D. Investigating the role of a large, shared display in Multi-Display Environments. *CSCW 20*, 6 (2011), 529–561.
35. Wilensky, U. and Resnick, M. Thinking in levels: A dynamic systems approach to making sense of the world. *J. of Science Education and Technology 8*, 1 (1999), 3–19.
36. Woodgate, D., Fraser, D.S., Gower, A., et al. Using mobile and pervasive technologies to engage formal and informal learners in scientific debate. In T.T. Goh, ed., *Multiplatform E-Learning Systems and Technologies: Mobile Devices for Ubiquitous ICT-Based Education*. Information Science Reference, Hershey, PA, 2010, 196–214.