Designing for Student Interactions: The Role of Embodied Interactions in Mediating Collective Inquiry in an Immersive Simulation

Michelle Lui
University of Toronto & University of Illinois at Chicago
michelle.lui@utoronto.ca, mmylui@uic.edu

ABSTRACT
Advances in mobile and wireless technologies provide new possibilities for supporting K-12 learning activities that can be spatially distributed in the classroom, for example in jointly investigating a scientific phenomenon. Such technologies have an impact on the ways in which students engage with one another, and with the quality of their engagement with the activity itself. This paper uses an embodied approach to understand the patterns of interactions between students (e.g., student-to-student, student-to-teacher) and with computational media within the environment (e.g., student-to-device, student-to-large display), in relation to students’ real-time meaning making as they engage in collective inquiry in an immersive simulation environment. The design-based research study consists of two iterations tested in an authentic school setting. We found that increased student-to-student interactions was accompanied by improved observational accuracy and higher quality student explanations constructed. The design implications of the research findings are discussed.

Author Keywords
Digitally augmented physical spaces; science inquiry; large displays; multi-device environments; embodiment; multimodality.

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

INTRODUCTION
Mobile and wireless technologies bring new possibilities for supporting co-located collaborative learning activities in K-12 classrooms. While computer-supported collaborative learning (CSCL) is already an established field of study, new technologies expand the design space of activity structures available to students and teachers [17,24]. In scientific inquiry, an 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., [3,7,8]). The primary goals of many technology-enhanced environments aim to support inquiry, helping students develop deep understanding of complicated concepts while also attaining scientific reasoning, critical thinking skills, and collaboration and communication skills [18].

In recent years, researchers and instructional designers have been taking advantage of handheld computers interconnected by a wireless network. This allows learners to work on a common task while interacting face-to-face, maintaining the mediation afforded by a technology-based system [30]. Coupling networked handhelds with spatially distributed displays (e.g., interactive whiteboards) and tabletop interaction systems can situate students in a number of different learning contexts, offering new ways to engage students with science concepts that have traditionally been taught or addressed through more passive forms of instruction. There is a growing body of literature on digitally augmented activity spaces (e.g., mixed reality environments) ranging from participatory simulations that enable students to jointly experience computational simulations of scientific phenomena as a honeybee in their system [21] to those more rooted in investigation practices that allow students to take on roles of domain scientists like seismologists as they investigate an earthquake occurring within their classroom [17], to full-body embodied activities that engage elementary students in play-acting how water molecules would behave in different environmental conditions, such as a cold winter day, in order to understand particles dynamics [2]. While small group collaboration and peer interaction form the basis for many of these complex activities, few have examined broad patterns of large groups of students interacting in such spaces.
BACKGROUND

Digitally Augmented Physical Spaces for Learning

Digitally augmented activity spaces are generally designed to support group activity in fairly structured settings in which participants are focused on a shared goal during a fixed time period, thus providing important experience-based learning opportunities that are socially and culturally situated [22]. We are just now starting to understand that how students engage with devices and displays in digitally augmented physical spaces have effects on how students engage with each other and their own thinking. Research efforts on collaborative activities in classroom settings using different devices such as handhelds, interactive whiteboards, and tabletop interaction systems suggest emerging patterns of interaction and classroom dynamics that may support learning in many ways [4,12,30]. For example, Liu and Kao demonstrated that shared display groupware promoted shared understanding of the workspace and increased awareness of partner actions [12].

By moving digital activity away from stationary modes of computing, mobile technologies have an impact on the production of social and embodied space through engaged interaction and emerging practices [22]. In Dourish’s seminal text, he defines embodied interaction as “the creation, manipulation, and sharing of meaning through engaged interaction with artifacts” (p. 126) [5]. His understanding of embodiment as “possessing and acting through a physical manifestation in the world” (p. 100) [5] raises the notion that the physicality of the environment can influence the interactions amongst peers and how they construct shared meaning. Some mixed reality learning environments examine embodiment in terms of the congruency between actions (e.g., gestures, full body movements) and concepts (e.g., gravitational force) [9]. Within our context of engaging students in inquiry as a collective (or learning community), we focus on embodied interactions as a means to understand learners’ relationships to 1) the various technologies around them, and 2) as communicative forms (e.g., position, gaze, gesture) for relating to other learners in the co-located space. In an effort to understand how to design complex learning environments, including learning materials that mediate collaborative knowledge construction, this paper examines student interactions with resources and with their peers as they engage as individuals and in small groups during a collective inquiry activity.

Small Group Interactions to Embodied Interactions

There is a body of literature on peer interaction and learning with computers in small groups that recognizes the potential for students to learn from each other, rather than only from their teacher [26,27]. Recent work by Price and Jewitt, offers a methodological approach to examine embodiment with case study analysis that draws on multimodal methods of analysis [23]. Multimodal analysis approaches are based on the premise that human communication involves both spoken and written language as well as non-verbal behavior that inherently form notions of embodiment, such as body movement and gaze [19]. That is, our interaction with the material world is characterized by communicative modes. By breaking down enactments of complex learning environments into units of action—mediated or otherwise, multimodality offers a way to explore concepts of embodiment in the knowledge construction process.

As with analyzing small group interactions and video-based interaction data, one challenge of utilizing multimodal methods of analysis is that they are often nuanced and specific to the context being examined. With increasing levels of complexity designed into mixed reality learning environments, detailed scrutiny of each and every student’s modes of communication for the entire duration of the intervention may not be feasible. Content analysis of qualitative data [1] may be used in such cases to interpret wider patterns of activity, as exemplified by CORTRA diagrams [6] and coding representations that depict user interactions over time [15,28]. For mixed reality learning environments and other complex spaces, quantitative data analysis offers broad brush sketches of embodied interactions in situ, while deeper analyses of small group interaction and discourse data allow us to understand how observed patterns translate into student actions, and how knowledge construction and collaboration is mediated. This paper uses multimodality as a lens to explore how the mixed reality environment can serve to support students in their collaborations and in using their collective knowledge to solve an inquiry problem. Specifically, this paper considers the role of embodied interactions in mediating their collective inquiry.

EVOROOM

Immersive Simulation Environment

EvoRoom is a collective immersive simulation that situates students in the rainforests of Borneo in Southeast Asia through wall-sized projected displays (“immersive walls”) that students examine during inquiry activities (Figure 1).

Figure 1. The EvoRoom environment, with projected wall displays, tablet devices, and collective visualizations.

Students take on the role of “field researchers” to complete tasks delivered to them on their personal tablet computers (“devices”), working in various group configurations as part of the inquiry activity. The tablets place students in small
groups, scaffold their activities, collect observations, and give real-time updates and individualized resources. Student work is aggregated and displayed on interactive whiteboards located at the front of the room to show real-time visualizations of collective work ("collective visualizations"). The light is dimmed and a soundtrack of ambient rainforest sounds serves to improve the sensory immersion of the students’ experience.

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 multiple weeks, the integrated curriculum included in-class activities, homework, a field trip to the zoo, and two collective inquiry activities with immersive simulations.

The first collective inquiry activity focused on evolutionary concepts. Students gathered evidence of evolution by observing changes in life forms as the simulation advanced through time. Findings from the pilot study and the first design iteration of this activity were previously reported. These early studies found EvoRoom 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 second collective inquiry activity was on the topic of biodiversity. Students predicted how an environmental factor (e.g., tsunami, earthquake, low rainfall) could change the biodiversity in Borneo’s ecosystem, and were challenged to explore the differences between four versions of the rainforest in order to locate the version that resulted from the variable they explored in their predictions. A previous report on the biodiversity activity examined how learners collected and explored data from the large displays, with particular emphasis on how students engaged with the emergent collective visualizations [14].

The current paper focuses on exploring multimodality as a methodological approach to examine embodiment and learning in a pilot study and two design iterations of the evolutionary activity. We give a brief report of our pilot study findings to set the context for the design choices made in the following two iterations. This work builds on previous reports on learning by examining students’ embodied interaction data with respect to their learning outcomes.

Research Questions
1. How are students’ interactions with each other and their teacher mediated by computational media in an immersive simulation during collective inquiry?
2. How can embodied interactions enacted within an immersive simulation support collaborative knowledge construction?

METHOD

Co-design
Using a co-design methodology [20] to design the simulation, inquiry activity and interactive materials, our core design team consisted of a graduate student researcher, a high school biology teacher, and a senior researcher. Over the course of our two-and-a-half-year collaboration, we met approximately once per week during the academic school year to consider design elements and outline our curricular strategy. In the months leading up to the enactments, design meetings widened to include two technology developers and an additional researcher in the second iteration.

Participants
This study was conducted at a university preparatory school in Ontario, Canada. The pilot study was conducted with eight female volunteers (ages 15-17) who had previously completed a Grade 11 Biology course. For the two iterations that followed, our co-design teacher's Grade 11 Biology students participated in the study (two class sections per year; ages 14-16). 45 students (25 male, 20 female) participated in the first iteration, with 56 students (26 male, 30 female) the following year.

Procedure
The pilot was a standalone activity, while the collective inquiry activities in iterations 1 and 2 were designed to fit within an integrated curriculum on biodiversity and evolution topics that spanned 10 to 12-weeks. The present paper focuses on the same collective inquiry activity across all three enactments within the EvoRoom environment. The collective immersive simulation was set up in a small classroom at our partner high school, with space to accommodate up to 16 students in the room at a time. As a result, each class section was split into two "sessions" (with a total of four per year of enactment). Following a design-based research method, iterative improvements were made to each design, all of which follows the broad phases of activity below:

1. Individual observation / data collection of species at various time periods (which were aggregated to collective visualizations in iterations 1 and 2)
2. Group reflection about the environment (where students synthesized information)
3. Evaluation of collective information and scientific explanations (which were again aggregated to collective visualizations)

Data sources
During the activity, video recordings captured student interactions, providing a data source for qualitative analysis, while artifacts created by students (e.g., observations, reflective notes, explanations) were also collected as measures of student performance. For iterations 1 and 2, video cameras were positioned in fixed locations around the room for recording different views (Figure 2) of participants interacting with the various components of the room and with each other.

Analysis
To explore embodied interactions during collective inquiry in the immersive simulation environment, we examined
patterns of changes in communications modes over time. We determined the relevance of established communicative modes in our mixed reality environment based on the multimodal interaction analysis framework [19]. From this, we generated a list of codes for examining gaze and interactions as applied to the environment.

Figure 2. Video camera positions (A-E) and designated spaces (1-6) in EvoRoom, in iterations 1 and 2.

<table>
<thead>
<tr>
<th>Communicative mode</th>
<th>Analyzed in EvoRoom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spoken language</td>
<td>Yes, coded as: student-student interaction, student-teacher interaction, teacher-student interaction, teacher-small-group interaction, and teacher-whole-group interaction</td>
</tr>
<tr>
<td>Proxemics</td>
<td>Not coded, but examined in detailed qualitative case analysis</td>
</tr>
<tr>
<td>Posture</td>
<td>Not coded, but examined in detailed qualitative case analysis</td>
</tr>
<tr>
<td>Gesture</td>
<td>Yes, coded as one of four types [16]: iconic gestures (e.g., mimicking action), metaphoric gestures (e.g., expression of abstract idea), deictic gestures (e.g., pointing), beat gestures (e.g., rhythm in speech)</td>
</tr>
<tr>
<td>Head movement</td>
<td>Not coded, but examined in detailed qualitative case analysis</td>
</tr>
<tr>
<td>Gaze</td>
<td>Yes, coded in terms of potential gaze targets: device, immersive wall display, collective visualization, student, teacher</td>
</tr>
<tr>
<td>Print</td>
<td>Yes, coded as: student-device interaction, student-visualization interaction, teacher-device interaction, teacher-visualization interaction</td>
</tr>
</tbody>
</table>

Table 1. Multimodal interaction analysis framework (modified from [19]).

Video data from various camera recordings from one session (in each iteration) were composited into one file, and then split into 2-minute segments. This gave us a view of almost everyone in the room at any given time. The first 10-seconds of each segment was analyzed for the interactions from Table 1.

Student artifacts created during the activity offered measures of task performance and the knowledge construction progress. Student observations in iterations 1 and 2 were analyzed for accuracy, while the explanations themselves were analyzed using a 0-5 Knowledge Integration scale that rewards valid scientific connections between concepts (Table 2).

<table>
<thead>
<tr>
<th>Score</th>
<th>KI Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No answer</td>
<td>Response is off task or “I don’t know”; Student writes some text, but it does not answer the question being asked</td>
</tr>
<tr>
<td>1</td>
<td>Off task</td>
<td>Have relevant ideas but fail to recognize links between them; Have irrelevant ideas</td>
</tr>
<tr>
<td>2</td>
<td>Irrelevant/Incorrect</td>
<td>Make links between relevant and irrelevant ideas; Have incorrect/irrelevant ideas</td>
</tr>
<tr>
<td>3</td>
<td>Partial</td>
<td>Have relevant ideas but do not fully elaborate links between them in a given context</td>
</tr>
<tr>
<td>4</td>
<td>Basic</td>
<td>Elaborate a scientifically valid link between two ideas relevant to a given context</td>
</tr>
<tr>
<td>5</td>
<td>Complex</td>
<td>Elaborate two or more scientifically valid links among ideas relevant to a given context</td>
</tr>
</tbody>
</table>

Table 2. Knowledge Integration rubric used to score student explanations (modified from [10]).

PILOT STUDY DESCRIPTION

The pilot was designed as a single 90-minute session, which took part one week after the end of the academic school year. After the premise of the activity was introduced, students individually recorded open-ended observations about various species in the rainforest environment at 2 million years ago (MYA; the area at the time was known as Sundaland). Once completed, the teacher advanced the room to ‘present day’ using her tablet device (that controlled various elements of the room), revealing a sequence of geologic events that affected the landscape over the span of two million years. The central landmass broke into a peninsula and several islands, including Borneo and Sumatra. While one side of the room showed Borneo’s ecosystem, the other side showed Sumatra’s. Students again made open-ended observations about the species on the opposing sides of the room.

Next, in the group reflection phase, students were divided into two groups of four students, with one group specializing
on Borneo and another on Sumatra. Each group answered a set of questions designed for students to review and compared notes about their individual observations (e.g., in the Borneo group, students were asked *What common species were found in both Sundaland and Borneo?*).

The final explanation phase asked students to choose organisms from the Borneo ecosystem and explain processes of evolution (e.g., adaptation) that may have occurred over two million years. The explanations were aggregated to the interactive whiteboard that the teacher used to lead a synthesis discussion to close the activity.

**PILOT STUDY RESULTS**

**Student Interactions**

Relatively few student-to-student interactions were observed (with the exception of the group reflection phase, where students were instructed to complete a collaborative task). Since fixed cameras were not yet implemented, only observed instances of peer, teacher, and device interactions were recorded. In the observation phase, which took place in approximately 30 minutes (of the total 90-min activity) 17 10-second video clips were analyzed. We observed student-to-student interactions in 12% of the video segments, and student-teacher interactions in 12% of video segments. This is likely due in part to the students remaining seated and stationary throughout, with access only to their immediate neighbours.

**Activity Performance**

Using a note-taking form on their tablet devices, students made free-form observations about the organisms shown in the simulation. A total of 157 observations were made and analyzed following Chi’s method for content analysis [1]. There was a wide range in the quality of the observations. For example, the word count ranged between 1 and 104 words per observation note. Many observations were rather basic and did not reveal information pertaining to evolutionary change. Moreover, over half of the observations made focused on the behavior of the organisms, which was not intended as a productive line of reasoning in the design (i.e., the behaviours of species were not designed so as to provide any productive evidence about evolution). For example, an observation that focused on behaviour was “There are two tapirs, one walking really slowly and one drinking from a shallow pool.”

14 explanations collected were scored a KI average of 2.36 (SD=0.75) out of a possible score of 5, which reflected the nature of the explanations, as many were about surface features of the species observed.

**DESIGN ITERATION 1**

A key finding from the pilot was that the enactment was characterized by limited interactions amongst stationary students. Early on, we took for granted that students would automatically move around the room simply because there was room to move and that they would talk to one another as they work simply because they were allowed to do so. When the students first walked into the room, they moved around in the space, looking around the room until they were asked to take a seat on one of the stools provided in the center of the room. The intent was for students to sit while the teacher introduced them to the activity, which they did. However, they remained seated there, looking around the room but did not leave their stools throughout the observation phase. Despite the teacher’s suggestion that they could see more if they went up closer to the displays, students remained in their seats, occasionally turning their bodies to switch views between large projected displays on opposite sides of the room. One student stood up briefly to look at one of the walls and sat down again. One idea to explore further was compelling students to move around the room, encouraging more natural face-to-face interactions amongst peers. It was one of our initial suppositions for such a space—that face-to-face interactions situated in an embodied physical space foster knowledge co-construction, thereby improving scientific reasoning and understanding.

**Design change:** Encourage student interactions through movement around the physical space of the immersive simulation by designing activities and prompts that required student localization.

**ITERATION 1 DESCRIPTION**

To provide a richer context for scientific investigation, we greatly extended the timeline from two (2 MYA and present day) to span 200 million years. In the approximately 115-minute-long activity, students examined the rainforest as it may have appeared at nine different time periods: 200, 150, 100, 50, 25, 10, 5, 2 million years ago and present-day Borneo and Sumatra. To implement student localization, content for each time period was indexed to different areas of the room, and students were directed around the room using QR codes to register their presence. For the observation phase, students were asked to look for assigned species. If the species were not present, they were asked to identify their predecessors from a short list and their accompanying descriptions (e.g., *Which of the following is most likely your organism's ancestor?*). Answers were recorded, resulting in the emergence of an aggregated cladogram on the interactive whiteboards.

For the group reflection phase, students were assigned to teams of 4-6. Each student was asked to discuss a specific question with his or her team and record the answer on their tablet (e.g., *Primates differentiated at 50 MYA. What did you see happen in 100 MYA in other categories that could explain this differentiation?*). The students’ answers were shown at the front interactive whiteboards.

In the explanation phase, teams used the aggregated information presented on the interactive whiteboards, including the cladogram, to articulate the processes of evolution for different species over the 200-million-year span. Each student was instructed to discuss with their team, but was ultimately responsible for contributing explanations about his or her assigned organisms.
ITERATION 1 RESULTS

Student Interactions
A pattern of frequent student-device, student-student, and student-teacher interactions, was observed. In the 57 10-second video clips analyzed, at least one student of ten (3 female, 7 male) was looking at their device (Figure 3). In 63% of the video segments, a student was observed to be looking at a fellow student, followed by gazes aimed at the collective visualizations (54%) and at the teacher (35%). Students interacted with devices the most (84%) of video clips, followed by other students (63%) and interactions with their teacher (61%). Inter-rater reliability (IRR) was performed on 20% of data, where substantial agreement (Kappa=0.91, p<0.001) was achieved.

![Figure 3. Interaction patterns the EvoRoom evolution activity session in iteration 1.](image)

Activity Performance
Of 1112 observations collected via student tablets across all four sessions, 78% were correct, with an upward trend of accuracy compared to time period (Figure 4). Since the evolutionary lineages of most organisms become less ambiguous as the time periods reach closer to present-day, we believe that the observation accuracy trend indicates that students were engaging with the media appropriately.

![Figure 4. Observation accuracy by time period in EvoRoom evolution activity in iterations 1 and 2.](image)

43 final explanations were collected, with an average word count of 33.28 (SD=29.51), and average KI score of 2.72 (SD=1.05). While explanations from iteration 1 were scored higher (M=2.72, SD=1.05) than those from the pilot (M=2.36, SD=0.75), the difference was not significant. In general, there was an increase in the complexity and sophistication of explanations from the pilot (34%) to those in iteration 1 (43%).

DESIGN ITERATION 2
In iteration 1, students were interacting within EvoRoom as initially envisioned—moving around, and talking to one another and referring to walls and tablets. However, related to the finding that students looked directly at another student in only 63% of video clips, the most prevalent kinds of student interactions can be described as students looking at or typing on their tablets, or looking at the immersive walls during conversations. In addition, during the observation phase, students were seen to move around, following instructions in completing the straightforward tasks, which did not did not optimize benefits of co-located peers for collaborative discussions about these observations. The findings suggest that more focus needed to be paid to improving the quality of student-to-student interactions.

Design change: Offer more opportunities for student-student interactions prompting engagement with other students where the media content serves as a source of discussion or cognitive support.

ITERATION 2 DESCRIPTION
In order to maximize the limited time during the activity for more meaningful interactions, we removed the stringent QR code scanning, while maintaining the indexicality of the content to specific areas. Students were either directed by their tablets or verbally by their teacher to different locations. In addition, we offered more opportunities for exchanges in the observation phase. Certain students acted as “guides” who were provided with additional information about designated time periods to support individuals with their observations. Similar to a docent’s role in a museum, students could receive guidance about their assigned organism’s presence and potential predecessors.

In the group reflection phase, students worked in groups of four to five, and compared adjacent time periods (e.g., 200 vs. 150 MYA) with reflective question prompts (e.g., Which species appeared in this time period that wasn’t there before? Consider climate, habitat, animals, and plants). Each student was assigned functional roles. For each question, one person was assigned to be the scribe, while others were provided with additional information about specific time periods on their tablets that served to augment content from the simulations.

The final step where students discussed and posted ideas about evolutionary processes was analogous to the question from iteration 1, but was further structured to ask students to provide evidence for their thinking. There was also explicit instruction to review the collective cladogram as part of their reflective process.
ITERATION 2 RESULTS

Student Interactions

Two sessions of the evolution activity were examined, each with a different condition for supporting the activity. While one session with 13 students (8 male, 5 female) used the custom tablet application (as per our original designs), another session with 14 students (5 male, 9 female) used paper materials (due to issues with network messaging technology). Aggregation of student work to collective visualizations could not be replicated with paper and were replaced by images of analogous representations (i.e., completed cladogram). The observation phase and group reflection phase were combined, which allowed students to examine each time period as a group.

A total of 31 and 29 10-second long video clips were analyzed at 2-minute intervals in the paper and tablet sessions, respectively. IRR was performed on 20% of data, where substantial agreement (Kappa=0.81, p<0.001) was achieved. Students interacted with other students frequently in both paper and tablet sessions of the activity, with slightly more student-to-student interactions observed in the tablet session (90% of video clips) compared to the paper session (84% of video clips; Figures 5 and 6). In the tablet session, this was followed by student-device interaction (76% of video clips). There were fewer occurrences of student-device interaction in the paper session—the equivalent of student-device interaction translated to writing on paper handouts. This was expected, since fewer students had access to a “device” (i.e., there were only two sets of paper handouts per group, a set of questions and a set of paper resources with additional information, paralleling what was available on the tablets), but it was still the second most frequent type of student interaction (55% of video clips). Student-teacher interactions were also observed in 55% of video clips in the paper session and 48% from the tablet session. There were more incidences of student-targeted gazes and student interactions in both sessions compared to iteration 1 (Figure 7).

In the pilot study, we observed an average of 0.12 student-student interactions (per video segment) during the observation phase, as compared to 0.94 and 3.12 in iterations 1 and 2 respectively (Figure 8), with student-device interactions remaining at relatively consistent levels.
Activity Performance

Students also achieved higher observational accuracy scores than in the previous year. 335 answers to the question: Which of the following is most likely your organism’s ancestor? were collected by the tablets, of which 84% were accurate across all time periods. As in iteration 1, there was an upward trend of accuracy compared to time period (presumably because the organisms shared more similarities with their present-day counterparts in more recent times). In iteration 2, over 90% accuracy was reached by 100 MYA, whereas in the previous iteration, over 90% accuracy was reached in two time periods (25 and 2 MYA; Figure 4).

At the end of the activity in all three versions of the evolution activity (i.e., pilot, iteration 1 and 2), students reflected on the evolutionary forces that were at play in the environment through 200 million years, allowing the culminating explanations to be compared. Due to time and technological constraints, students in the tablet group were unable to complete the final step of discussing and posting ideas about evolutionary processes. This was remedied by asking students to complete the assignment as homework and discussing as a class in the following period.

38 explanations were written and the answers were coded using the KI scale, revealing that students who participated in the iteration 2’s tablet and paper sessions achieved a mean KI score of 3.7 (M=3.71, SD=1.14), showing successively improvement over the pilot and iteration 1. Knowledge Integration scores for students’ explanation of evolutionary processes in iteration 2 were significantly higher than those in iteration 1 and the pilot study, F(2, 92) = 12.64, p<0.001 (Figure 9).

![Figure 9. Mean KI scores of evolution explanations in pilot and iterations 1 and 2 of EvoRoom evolution activity.](image)

Correspondingly, there were also more explanations that were coded as “complex,” as over 50% of the notes were coded either as Basic or Complex (Figure 10). The improvement in iteration 2 explanation scores is partly a result of improvement to the activity design, but may also have benefited from allowing students to complete the question after the activity. It could also be due to improved phrasing of the question, which included more scaffolding in asking students to review the cladogram.

![Figure 10. Distribution of evolution explanations’ KI scores in the pilot and iterations 1 and 2 of EvoRoom evolution activity.](image)

Multimodal Interaction Cases

Closer examination of student collaboration is revealed through multimodal interaction analysis, which showed distinct forms of collaboration between paper and tablet sessions. In the paper session, a group of four was seen to be standing while examining immersive walls (C, D) in order to make comparisons between two time periods. The group remained close together, but two members were more active participants of the discussion (Dan and Jody). Dan moved between walls C and D, frequently pointing or making pointing gestures towards the walls. Two less active members (Lucy and Jeremy) remained on the periphery, but observing the walls and paying attention to the discussion. She offered one piece of information about the tapir, but otherwise kept observing with her arms crossed. The fourth member, Jeremy, did not contribute but remained close to the group, except when walking between walls C and D to observe the content more carefully.

The somewhat uneven patterns of participation in this group interaction reflect challenges that are common to many classrooms, where some students participate more actively than others. In the case of this particular group, the two less active members seemed to be thoughtful about the content of the simulations, either making well-timed comments (indicating that they were listening to the exchange while looking at the large displays) or moving between “time periods” to make comparisons. It appears that this group of students could have benefited from collaboration scaffolds that assigned roles to each student and directed members’ attention to specific information. Since functional roles were not distributed evenly in the paper session, this led to some members not participating fully. Although the students were observing walls content alongside their more active counterparts (who tended to hold the set of paper resources), they did not make many verbal contributions to group discussion. This also placed the burden of responsibility in completing the work on the person who held the pen.
In the group reflection phase of the tablet session, two of the three groups sat down in their designated corners of the room (as directed by their tablets and their teacher) to answer their assigned questions while a third group stood in front of the walls during part of their discussion. A multimodal interaction analysis of the third group revealed how they worked together in the presence of collaborations scaffolds provided by the tablet devices (i.e., functional tasks assigned, screens limited to functions needed to support assigned tasks).

The group of four worked in pairs as well as a whole group. In the beginning, Phoebe and Sharon were observing the walls (A and B) while standing together, while Cory and Ray worked as a pair. After Phoebe turned towards Cory and Ray and asked the group, “What are the major differences are there?” gesturing towards the walls, the four members considered the differences together. Sharon shared information about a mass extinction that occurred prior to one of the time periods from her tablet, and when Phoebe had a question, Sharon pointed out information on Cory’s tablet. All four looked at his tablet as a common referent, but Ray was more interested in answering the question for which he was the scribe for, asking Cory for assistance. The teacher, in an attempt to help, inadvertently broke up the interaction by offering additional paper copies of resources (i.e., information about specific time periods).

While other groups worked collaboratively to answer the three assigned questions, one at a time, the members in this group chose to answer all three questions at the same time. Cory, Phoebe, and Ray were each responsible for answering a question relating to the comparison of the rainforests at 200 MYA and 150 MYA. Sharon was responsible for reviewing information about the area at 200 MYA, and once Cory submitted an answer to his question, was able to take on the information resource role for 150 MYA. Both resource roles supported scribe members during this interaction, but Ray and Phoebe had opposing goals: one sought to answer the question, and the other wanted to start a discussion. This interaction case showed that limiting information resources to a specific group member allowed them to support their scribe peers and also freed up the tablet screens of scribes. In fact, when the teacher showed Ray additional paper-based resources, he no longer needed to work with Cory and moved away from the group.

The group interactions examined in detail indicated that explanations were better constructed (based on their KI scores) in the tablet session compared to the paper session. This suggests that assigning localization, functional roles, and limiting resources relevant to specific roles, would result in an environment that is more conducive to collaborative reflection. Another example of functional roles being assigned to students in a different context was the use of guides in the observation phase of iteration 2.

**DISCUSSION**

This work revealed a number of opportunities for refining co-located investigative inquiry activities supported by technology, particularly concerning how to design collective knowledge co-construction experiences in digitally augmented physical spaces. In a mixed reality environment like EvoRoom, much of the appeal is related to the physical space within which the simulation is set—the “immersivity” of the room itself, and the opportunities it affords for naturalistic physical and social interactions. In considering which design improvements to make for each iteration, we attended to the quality of student interactions as well as in encouraging students in reflective practices and effective collaboration, which we believed helped to make progress on both fronts (student interactions, activity performance). Our findings speak to the kinds of activities that would be engaging and productive for learning within immersive simulations and related spaces. Below we highlight two key design considerations in addressing our research questions.

**Fostering interactions around shared referents**

Overall, our findings demonstrate increased student-to-student interactions through successive iterations, which highlighted design considerations of the immersive simulation as a shared referent (RQ1). In the pilot study, seated stationary students were focused in their interactions with their tablet device, as characterized by a head-down effect, with only brief exchanges amongst neighboring peers. To review immersive wall content, students looked up and back down again to type long-form observations, occasionally twisting their bodies and repositioning themselves from their seats to look from one side of the room to the next. In the group reflection phase, students tended to rely on their own observations, and students were only able to make superficial evolutionary connections. The immersive walls and aggregated displays thus acted as asynchronous sources of information rather than shared referents. In iteration 1, we saw increased student-student interactions when content was indexed to distinct areas of the room and QR scanning required students to physically move around the room. However, more student gazes were directed towards their tablets than their co-located peers. In the group reflection phase, we observed longer discussions as they gazed at the immersive walls and collective cladogram, which demonstrated how the large displays were used as shared referents. In iteration 2, content localization and instructions were sufficient in directing student movement. With the addition of guides, there was increased student-student interactions as well as gazes directed at fellow peers. Although interactions between guide-participant pairs tended to be brief and consisted of question-and-answer exchanges; still, the addition of guides allowed students to quickly move through a straightforward task, and set the stage for peer interactions involving specialized roles throughout the activity. Students also achieved higher observational accuracy scores than in the previous year.
Increased student-student interactions through successive design iterations were accompanied with improvements to their activity performance (i.e., observations, explanations). While we are not suggesting that increased frequency of student interactions and their performance improvements are directly correlated, we believe the different modes of interaction examined (e.g., gaze, gesture) reflect the extent to which notions of embodiment either mediated or contributed to students’ meaning making. Related literature on patterns of interaction of computer-supported learning environments suggests high quality interactions between students may be characterized by sustained interactions and on-topic discourse [11]. Other studies on multimodal interactions and learning (e.g., [23,25]) report that high quality interactions are marked by the presence of gestures (indicating engagement in discourse), attention towards one another (e.g., gaze), and posture (consider a pair of students looking at other components with their shoulders directed away from one another, compared with a pair looking at each other, their shoulders directed at each other, using gestures to denote external information into their discourse).

Interdependency in collaborative work
A means of designing interdependency among peers was through task-based (i.e., functional) roles within the activity. This was best demonstrated in the tablet session of iteration 2, when students were provided with collaboration scaffolds in the group reflection phase. From their tablets, students were asked to designate a person to act as the scribe of the group (for a particular question), as well time period specialists (i.e., responsible for looking up resources about the rainforest at each assigned time periods).

When group members’ tablets presented different information (e.g., person tasked with looking up information at 200 MYA vs. 150 MYA had access to facts about the corresponding time periods), the tablet was observed to be a mediating object among group members. Students referred to the tablet as a source of information within an interaction. As a point of comparison, students in the pilot were asked to collaborate, but the materials during group reflection amongst group members were identical. In iteration 1, each student was assigned different species for which they were responsible for in their group discussion, but student interactions in both enactments did not engage all members fully.

Our insights about the importance of designing interdependency into the activity and resource allocation echo the findings from recent studies about pairs of students engaging in a tabletop game, which found that those in the roles condition more frequently gave explanations that moved their line of inquiry forward [29]. This paper builds onto their work by demonstrating distinct patterns of interaction in collaborative activities that relied on interdependent roles, providing multimodal sketches of students’ experiences under these conditions.

The patterns of student interaction data and multimodal interaction cases analyzed in iteration 2, which demonstrated higher levels student-student interaction, increased gazes towards other students, and supportive local group collaboration between scribe and information resource roles, highlight the relationship between instructional design decisions (e.g., task-based specialization, allocation of resources) and students’ embodied interactions in how they related to one another (RQ2).

Limitations
One limitation of this study concerns the method of analysis. For each video clip, the presence of the interactions was noted for each student. While this allowed some analysis of the kinds of interactions that occurred throughout the activity, the method of splitting video data into 2-minute segments and sampling only the first 10-seconds likely introduced a risk of over representation for short-duration interactions. Nevertheless, this approach did allow some unbiased inspection of basic forms of interactions that occurred at various phases during the activities, which is useful for comparison between activities (and iterations), and allows identification of interesting patterns of interactions for deeper analysis (i.e., using a multimodal interaction analysis method [19,23]).

CONCLUSIONS
This paper explored broader patterns of multimodal communication as an accompanying analysis to traditional multimodal interaction analysis. This approach was applied to an investigation on the role of embodied interaction in mediating collective inquiry processes in mixed reality environments, wherein representative patterns of interaction and dominant modes of communications in the environment were revealed. We found that increased student-to-student interactions was accompanied by improved observational accuracy and higher quality student explanations.

ACKNOWLEDGMENTS
Special thanks to my advisor, Jim Slotta, for his support and valuable insight throughout the project, to María Niño-Soto for her extensive contributions as our design partner. Thank you to members of Encore Lab, who contributed to the development and data collection effort, especially Alisa Acosta, Colin McCann, Armin Krauss, Matt Zukowski, and Gabby Resch. I am indebted to the many student volunteers who have participated in this project. This material is based on work was supported in part by the Social Sciences and Humanities Research Council of Canada Grant 410-2008-2125.

REFERENCES
2. Joshua A Danish, Noel Enyedy, Asmalina Saleh, Christine Lee, and Alejandro Andrade. 2015. Science through Technology Enhanced Play: Designing to support reflection through play and


