Bridging Physical and Virtual Learning: A Mixed-Reality System for Early Science

Nesra Yannier

Ph.D. Thesis Proposal, November 16, 2015
Human Computer Interaction Institute

Committee:

Ken Koedinger (Chair) (HCI & Psychology, CMU)

Scott Hudson (Chair) (HCI, CMU)

Jessica Hammer (HCI & Entertainment Technology Center, CMU)

Kevin Crowley (Learning, Research and Development Center, University of Pittsburgh)

Abstract:

Tangible interfaces and mixed-reality environments have potential to bring together the advantages of physical and virtual environments to improve children's learning and enjoyment. However, there are too few controlled experiments that investigate whether experimenting with physical objects in the real world accompanied with interactive feedback may actually improve student learning compared to flat-screen interaction. Furthermore, we do not have a sufficient empirical basis for how a mixed-reality environment should be designed to maximize learning and enjoyment for children.

I created EarthShake, a mixed-reality game bridging physical and virtual worlds via Kinect depth-camera sensing to help children learn physics. I have conducted two controlled experiments with EarthShake that have identified features that are more and less important to student learning and enjoyment. The first experiment examined the effect of observing physical phenomena and collaboration (pairs versus solo), while the second experiment replicated the effect of observing physical phenomena while also testing whether adding simple physical control, such as shaking a tablet, improves learning and enjoyment. The experiments revealed that observing physical phenomena in the context of a mixed-reality game leads to significantly more learning (5 times more!) and enjoyment compared to equivalent screen-only versions, while adding simple physical control or changing group size (solo or pairs) do not have significant effects. Furthermore, gesture analysis provides insight as to why experiencing physical phenomena may enhance learning.

My proposed work will further investigate what features of a mixed-reality system yield better learning and enjoyment, especially in the context of limited experimental results from other mixed-reality learning research. Most mixed-reality environments, including tangible interfaces, currently emphasize open-ended exploration and problem solving, and are claimed to be most effective when used in a discovery-learning mode with minimal guidance. I plan to investigate how critical to learning and enjoyment interactive guidance and feedback is (e.g. predict/observe/explain prompting structure with interactive feedback), in the context of EarthShake. I propose to conduct an experiment that compares the learning and enjoyment outcomes of children interacting with a version of EarthShake that supports guided-discovery, another version that supports exploration in discovery-learning mode, and a version that is a combination of both guided-discovery and exploration. This thesis will also explore how this approach can be extended to other content areas, with the goal of creating a new mixed-reality system that can be used in museum and school settings to improve young children's science learning and enjoyment in a collaborative way, fostering productive dialogue and scientific curiosity.

Table of Contents:

1. INTRODUCTION

Today even very young children are being drawn into the compelling world of screen-based technologies, such as tablets or computer games. Screen-based technologies can help children learn by providing immediate targeted feedback (Corbett, Anderson, 2001). However, as screen-based technologies are becoming so appealing for children, it is worth asking whether real world interaction is still needed to enhance learning and enjoyment. Are today's children actually missing out on opportunities to develop understanding more readily in their physical environment by being immersed in these flat-screen technologies? In fact, some have argued that these flat-screen technologies can have negative effects on children (Roe, Mujis, 1998; Center for the Digital Future, 2009; Turkle, 2015).

How can games for young children combine the distinct advantages of screen-based and physical learning environments? In addition to instructional support, screen-based games also have potential motivational benefits, such as compelling scenarios and engaging characters. On the other hand, most natural learning occurs in our physical 3D world and some argue that that is where learning is at its best (e.g., Henning, 1998). The physical world has the potential to help children play, discover, experiment and learn in their everyday world and do so in a way that also supports social interaction (Turkle, 2015). Indeed, many past technology efforts have encouraged children to play with physical objects such as building blocks and puzzles to learn a variety of skills (O'Malley, Stanton-Fraser, 2004). Particularly in science domains, children's observations of changes in their everyday physical environment may aid them in more readily making discoveries and developing understanding of basic science principles. By combining the advantages of the physical environment and computer technologies, tangible interfaces and mixed-reality environments may help students learn in more engaging and powerful ways than either approach alone.

Although there are many compelling tangible interfaces, there are too few experimental tests of the hypothesis that physicality may improve student learning (Walker et al., 2012). Furthermore, we do not have a sufficient empirical basis for evaluating alternative explanations for why and how physicality may enhance learning. Most previous studies do not identify what it is that provides benefits for learning in these mixed-reality environments: does observing physical phenomena play an important role for learning in an interactive setting or is physical/hands-on control critical to enhance learning? For example, in ListenReader (Back et al., 2001), a paper-based book that has pages augmented with digital information does the hands-on action of turning pages provide any benefit? Or for BitBall (Resnick et al., 1998), is it more beneficial to observe a physical ball rather than a virtual ball on a flat-screen to learn the underlying principles of acceleration or does the action of throwing the ball provide any learning benefits? Through rigorous controlled experimentation, I hope to discover if physical experimentation and observation within an interactive flat-screen game can improve learning and enjoyment. I further hope to understand under what circumstances physical interaction will provide benefit. My review of the literature suggests that this thesis presents *the first* randomized tightly-controlled experiments establishing that physical observation in the context of an interactive mixed-reality game can improve engagement and learning for children above and beyond that produced by a matched flat-screen (non physical) control.

The following sections describe prior work and provide theoretical background. First, I discuss the mixed results from education research comparing learning with physical materials to learning with flat-screen analogs. While this work shows benefits for physical over virtual interactions in some cases, it mostly demonstrates how little we know about what makes 3D interactions useful, and under what conditions. Next I review work on everyday objects that have been instrumented with technology: *tangible interfaces* and *mixed-reality environments*. The range of work in this area shows how many technical challenges have been overcome in integrating computation with physical objects for learning. However, the literature also reveals a lack of experiments that measure learning with these interfaces, especially when compared with rigorous controls (Walker et al., 2012). This thesis is an attempt to begin to answer the questions left by both literatures.

1.1 Learning with Physical Objects

Experiments on the role of physical objects in learning have produced mixed results. I first present research that found benefits for physical objects over 2D representations of the same concepts. Hayne et al. demonstrated that 2 and 3 year olds can learn the assembly of a simple toy quite easily from watching a person, but have difficulty learning from a video of that person (Hayne et al., 2003). Martin and Schwartz showed that manipulating physical chips facilitated children's interpretation of fractions better than seeing a visualization of the grouped pieces on paper, though they only measured performance with these scaffolds, not learning after the scaffolds were removed (Martin, Schwartz, 2005). Other research has demonstrated no added learning benefit of physical materials over analogous virtual materials. Klahr et al. compared students' learning of experimental design principles when designing experiments with physical springs versus analogous virtual springs and found no difference for middle school students (Klahr et al., 2007). These experiments did not include any interactive feedback; students interacted with either physical or virtual materials on their own without receiving any instructional feedback on their actions. In another experiment in the context of light and color, Olympiou and Zacharias also found no difference in learning from only physical versus only virtual materials for university students. However, in the same experiment, they found that students who engaged in both physical and virtual interactions sequentially learned better than either the physical-only or virtual-only conditions (Olympiou, Zacharias, 2012).

These results suggest that there may be complementary benefits of learning from physical and virtual materials. Positive results appear to be more likely when physical and virtual environments are brought together or with younger participants. Such benefits may be further enhanced when physical and virtual materials are brought together in a mixed-reality environment, where children can experiment in their physical environment with interactive feedback. I aim to create a mixed-reality environment bringing together the advantages of physical and virtual environments to improve young children's science learning.

1.2 Mixed-Reality Environments and Tangible Interfaces for Learning

Mixed-reality environments and tangible interfaces bring together physical and virtual worlds by sensing physical interaction and providing output accordingly (Ullmer, Ishii, 2000). Mixed-reality learning environments can provide the benefits of physical objects while leveraging computational power to give students feedback and other instructional support. Many researchers have instrumented objects for learning to make them interactive, including, a book with an audio soundtrack that plays when the pages are turned (Back et al, 2001), a play-mat that records and plays stories (Ryokai, Cassell, 1999), a ball that measures and shows its acceleration (Resnick et al., 1998), a mixed reality experience that helps children discover and reflect on historical places and events (Stanton et al., 2003), and an interactive display for children to create, record, view, and test systems of tangible simple machine components (Tseng et al., 2011). Most research on mixed-reality learning environments has focused on studying them as prompts for student investigation and exploration and focusing on immediate effects on how students use these objects. While this body of work addresses the role of physicality in the immediate interaction, it does not well address the role of physicality in learning, that is, in long-term changes in how kids think. Such changes can only be reliably revealed in later assessments outside the tangible interface environment and are best evaluated in comparison to assessment outcomes that result from a reasonably alternative learning environment. For the most part, research in this area does not use post-test assessments of learning and does not include control conditions.

Instead of designing toward impact on immediate interaction, my goal has been to design toward improving student learning while also enhancing (or at least maintaining) student enjoyment. Given literature on benefits to learning of interactive forms of guidance, such as interactive feedback and self-explanation (Corbett, Anderson, 2001; Aleven, Koedinger, 2000), I designed a mixed reality learning environment that puts more emphasis on these supports and less on the kind of unguided exploration typical of much past research. My aim was to augment the physical environment with synchronized, interactive feedback and inquiry-based activities to produce a pedagogically strong and engaging learning experience that help children understand the reasons behind why things happen. Additionally, I wanted to determine the effects of physicality by using a post-test assessment to measure student learning, and by randomly assigning students to either a mixed-reality environment or a screen-only matched control.

Unlike the mixed results for non-instrumented physical objects, research comparing tangible and virtual interactions generally shows a benefit for tangibles (mostly performance benefits rather than learning outcomes with pre/post tests). Children were more successful and faster at solving puzzles when using tangible puzzle pieces instead of comparable interactions with a mouse (Antle et al, 2009).

Bakker et al. designed and evaluated MoSo Tangibles: a set of interactive, physical artifacts with which children can manipulate the pitch, volume and tempo of ongoing tones, in order to structure their understanding of these abstract sound concepts. Their results indicate that MoSo provided children with a physical handle to reason about targeted abstract concepts (with qualitative interviews and video analysis) (Bakker et al., 2011) Shelley et al. demonstrated problem solving and collaboration advantages for a paper-based tangible user interface for educational simulations over mouse interaction (Shelley et al., 2011). Logistic apprentices demonstrated enhanced task performance, collaborative interactions, and sense of playfulness when using a tangible instead of multi-touch interface (Schneider et al., 2011). In another study, students better remembered cause and effect relations in climate when they used a haptics-augmented environment where they could feel forces in addition to a virtual environment (Yannier et al., 2009). In another study, Yannier et al. created and evaluated FeelSleeve, an interface that allows children to feel story events in their hands via haptic feedback while they are reading on a mobile device. Their results showed that story events accompanied by haptic feedback are better comprehended and appear to be more salient in memory.

Although these studies provide support for the benefits of tangible interfaces and mixed-reality environments in education, we lack sufficient experimental research that tests whether these environments can produce learning benefits for children beyond simpler-to-develop flat-screen alternatives. Additionally, these studies do not identify *how* these environments benefit learners. Specifically, most of these studies confound two variables: observing phenomena in the physical environment and manipulating physical objects. To untangle the effects of each, we need randomized controlled experiments that isolate these variables.

1.3 Theoretical Background

Prior theoretical work offers several explanations for why observing changes in the physical environment in the context of a mixed-reality game may improve learning over an equivalent screen-based game: 1) *embodied cognition*: physicality facilitates mental visualizations and cues analogs to reason with; 2) *engagement*: physical experience is inherently more engaging; and 3) *collaboration:* physical environment provides more opportunities for collaboration which enhances learning. I discuss each in turn.

First of all, experiencing a physical phenomenon may help people perceive and mentally visualize the target objects (Englekamp, Zimmer, 1989; Abrahamson et al., 2014), leading to better understanding of scientific principles underlying physical phenomena. This mental visualization may then facilitate connections with familiar objects, and result in improved memory for the concepts related to those objects. Physical observations may be more deeply processed so as to recognize key features that explain physical phenomena (e.g. that a higher center of mass leads to instability). This theory follows Antle's research on embodied child-computer interaction, suggesting that when children (and adults) learn or reason with abstract concepts, they utilize mental simulations based on concrete motor-perceptual experiences (Antle, 2012). Also Hostetter et al. have theorized that perceptual and motor simulations underlie embodied language and mental imagery, which are often revealed by spontaneous gestures that accompany speech (Hostetter, Alibali, 2008). During a physical interaction, neural patterns of brain activity are formed across modalities. These patterns are integrated into a multimodal representation in memory. When such an experience is recalled, the multimodal representation is rerun, reactivating the same neural patterns (Antle, 2012). For example, repeated patterns of physically balancing the body give rise to neural patterns that are stored as a multimodal representation. This schema is activated when visually seeing balance and when thinking about balance in abstract domains such as mathematics (Abrahamson et al., 2012). Also, physical objects may trigger affordance for action, which in turn facilitates retrieval from memory. Research on embodiment shows that memory for actions (e.g. performing a command such as "open the book") is better than memory for the verbal description of the same commands (Glenberg, 1997). One interpretation is that memory specializes in embodied information. Thus, observing phenomena in the real world in a mixed-reality environment may trigger mental simulations and affordances for action, facilitating retrieval from memory.

Secondly, experiencing a physical phenomenon in real life may be inherently more engaging than watching a video of the same phenomenon, and thus be more powerful in directly supporting conceptual change. This claim is supported by Montessori's theory that young children are highly attracted to sensory development apparatus and that they use physical materials spontaneously, independently, and repeatedly with deep concentration (Montessori, 1964).

Finally, interacting in the physical environment may lead to more collaboration, which may in turn enhance learning. Shelley et al. have shown collaboration advantages of physicality (Shelley et al., 2011). Also, proponents of collaborative learning have claimed that the active exchange of ideas within small groups not only increases interest among the participants but also promotes critical thinking (Gokhale, 1995). Consequently, collaboration facilitated by physicality may improve learning.

Thus, adding physicality to an interactive game might improve learning for children. To test this hypothesis, I designed two carefully controlled experiments comparing learning outcomes within a simple interactive game with guided feedback.

In the first experiment, I compared the mixed-reality version of EarthShake (children observing physical phenomena with interactive feedback) with the virtual laptop version of the same game (where students watched videos of the same phenomenon integrated into otherwise equivalent screen-based version of the game). Additionally, to examine the effects of collaboration, within each game condition I compared students playing in pairs to students playing solo. In the second experiment, I again compared the mixed-reality versions of EarthShake with the equivalent screen-based versions. However this time I also added a potentially engaging simple physical control (such as shaking the tablet to create the earthquake on the screen) to investigate if adding an inherently more enjoyable physical/hands-on control can increase learning by increasing enjoyment or if physical observation and experimentation is more critical to enhance children's learning and enjoyment. Below I review EarthShake and the experiments in more detail.

2. EARTHSHAKE

Earthshake (see Fig. 1) is a mixed reality game that brings together the physical and virtual world to help children learn basic physics principles of stability and balance (Yannier et al., 2013). EarthShake aims to improve learning and social interaction by blending the advantages of computer games (engaging characters, compelling scenario, guided experimentation and immediate feedback) with the advantages of the physical environment (tangible learning, physical experimentation, discovery, and face-to-face social interaction and collaboration).

As shown in Fig 4, EarthShake consists of a multimodal interactive earthquake table, physical towers made of blocks, a Kinect depth camera and a display screen behind the table. It utilizes a predict/observe/explain cycle, where children are asked to make *predictions* about stability, *observe* outcomes of physical experiments, and *explain* those outcomes. The system detects which of the towers in the physical setup falls first when the user shakes the table and gives visual and audio feedback accordingly (Yannier et al., 2013). Children are guided by pedagogical prompts that highlight whether or not a prediction was correct and that scaffold explanations of the actual outcome.

The predict/observe/explain scaffolding sets a context in which children can construct an understanding of ideas such as symmetry and how they are relevant to physical properties of stability, consistent with theories of learning by doing and minimal assistance (e.g. Vygotsky, Dewey). Vygotsky argues that specific learning experiences can help people get from Zone of Proximal Development to the Independent zone; activities that help them assign different meaning to objects make them think independently (Vygotsky, 1978). Different structures that children reflect on while interacting with EarthShake may encourage them to think of blocks that they play with everyday in a different way, understanding the underlying physics principles that are relevant to their everyday experiences. Children are not directly told about the physics principles (symmetry, center of mass, wide base, height etc.) or how they are relevant (i.e. they are not told directly whether a tower is symmetrical or not and how that affects the tower's stability). They are able to discover these principles through real world feedback with pedagogical prompts on their predictions. To further facilitate mental construction of these key ideas, I use prompted self-explanation (Aleven, Koedinger, 2000).

EarthShake is targeted for children, ages four through eight (K-3rd grade) and aims to teach them principles of stability and balance, which are listed in the NRC Framework & Asset Science Curriculum for this age group (Quinn et al., 2012). It also builds on Azmitia and Crowley's research, which stresses the importance of scientific thinking and collaboration in an earthquake micro-world, specifically targeting principles such as wide base, height, symmetry, and center of mass, which are critical for understanding stability and balance (structures that are shorter, symmetrical, and have a wide base and lower center of mass tend to be more stable) (Azmitia, Crowley, 2001). It also utilizes contrasting cases, shown to be beneficial for deep understanding in science (Chase et al., 2010).

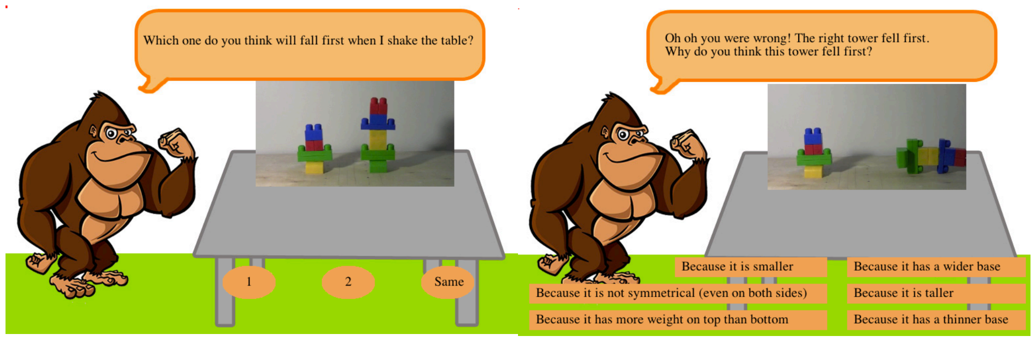

Figure 1. Virtual-only version of Earthshake, showing the predict/observe/explain cycle. The video of the physical towers shaking on the earthquake table is integrated into the game interface.

2.1 Scenario

Here I describe the mixed-reality version of EarthShake. EarthShake is structured around a predict/observe/explain cycle. The game starts with the gorilla character asking students which of the towers will fall first when he shakes the table (Yannier et al., 2013). The users can see prebuilt physical towers placed on a real earthquake table and, at the same time, a virtual representation of the same towers in a projected interface of the game behind the table. First, students use a mouse to click on the virtual representation of the tower that they predict will fall first. The gorilla then tells the users to discuss with their partner why they think this tower will fall first. When the students are done discussing, they click the "shake" button to shake the physical earthquake table and observe the results.

When the table shakes, the Kinect camera and computer vision algorithm determine which tower fell. If the students' prediction was correct, the gorilla says: "Good job! Your hypothesis was right. Why do you think this tower fell first?" If they were wrong, he says: "Oh oh you were wrong! Why do you think this tower fell first?" To explain why that tower fell, the students choose one of six explanations projected on the screen, providing scaffolding. The menu, read aloud by the gorilla, consists of the following choices: "Because it is smaller", "Because it is taller", "Because it has more weight on top than bottom", "Because it has a wider base", "Because it is not symmetrical", "Because it has a thinner base". (Figure 1) This scenario is repeated for different contrasting cases targeting height, wide base, symmetry, and center of mass principles (Figure 2). Note that while students observe the physical towers, they do not touch them.

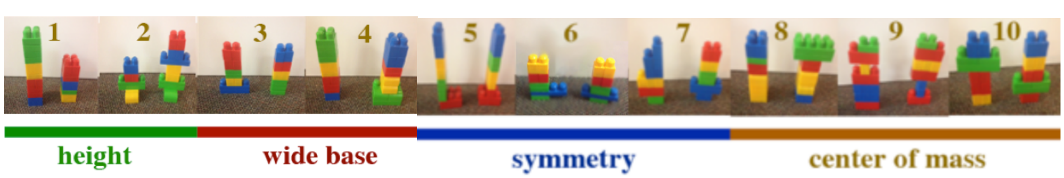

Figure 2. Contrasting cases used during the game.

2.2 Physical Setup and Vision Algorithm

The physical setup of EarthShake includes an earthquake table, physical towers placed on the table, a Kinect camera facing the tower, a projector, and a display screen with the computer game (Figure 3). The Kinect camera detects when a tower falls, ensuring that EarthShake is in sync with what is happening in the real world. The projected computer game provides visual and audio feedback to the user (e.g., noting which tower the student predicted would fall and which actually fell) (Yannier et al., 2013).

The earthquake table consists of a small motor, a switch, a disk and two layers of wood connected to rails. When the user pushes the switch, it activates the motor, which turns the disk, which then moves a rod connected to it. The rod is attached to the tabletop, which then moves back and forth.

The vision algorithm uses color segmentation and depth information to determine where the towers are located and to detect when they fall. Depth information reliably segregates the blocks from the background and eliminates conflicts that can arise when the background and blocks are similar colors. Simple blob tracking is then used to track each segment of the colored blocks. The size and location of these blobs are used to interpret the live state of the blocks on the screen. Finally, falls are detected when all blobs for a tower fall below a threshold height above the table (Fig. 4).

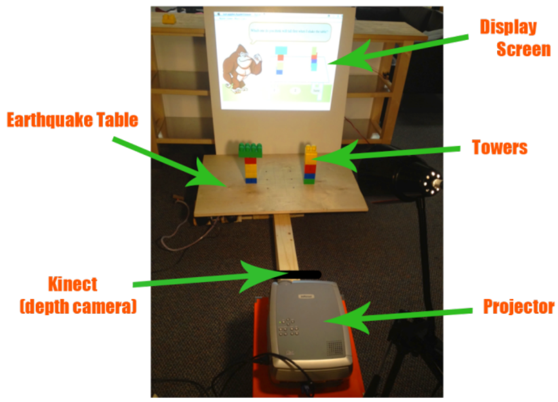

Figure 3. Physical setup of EarthShake.

From a technical perspective, the challenge is in creating tangible interfaces that are sophisticated enough to not only provide children with room for exploration, but also to provide them with interactive feedback that adapts to changes in the physical environment. Such feedback is critical for effective learning (Corbett, Anderson, 2001). Without technological support, it is often difficult in real-world tangible interaction to impose pedagogical structure and especially track students' actions. Such structure and logging is comparatively easy in purely virtual settings. I use the Kinect camera and a specialized vision algorithm to overcome this challenge.

Using Kinect to blend the physical and virtual environments also expands the paradigm of tangibility beyond specially instrumented objects. Many tangible systems require computation within the physical objects and are not affordable enough for widespread use. Systems such as MirageTable (Benko et al., 2012) and DuploTrack (Gupta et al., 2012) have demonstrated the potential of merging real and virtual worlds into a single spatial experience. With the introduction of inexpensive depth cameras such as the Microsoft Kinect, there is an opportunity for new, scalable paradigms for interaction with everyday physical objects.

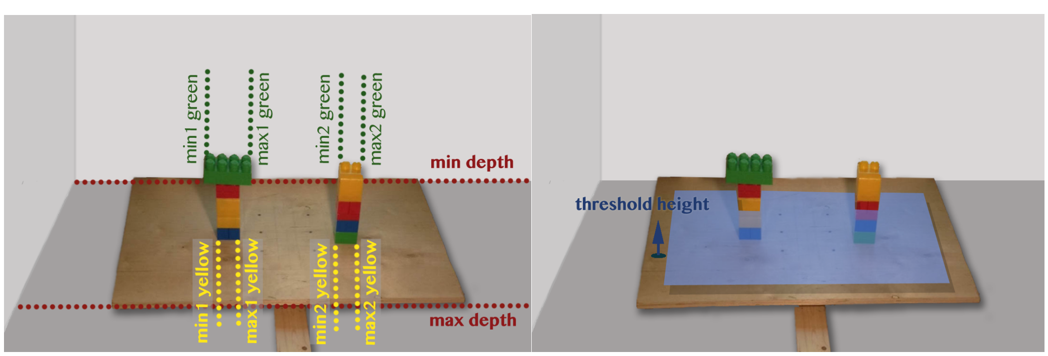

Figure 4. Specialized vision algorithm first calculates the minimum and maximum values for each color blob in each tower to determine where each tower stands (top). Then I use a threshold height to detect a fall, if all the color blobs in the tower are below this height (bottom).

3. PILOT STUDY

A single-condition pilot provided an initial evaluation of EarthShake's design and its effect on learning, usability, collaboration, and engagement. Twelve children participated (five female; grades K-3rd). The students played in three groups of two and one group of six, in a classroom setting. The study was conducted in a local elementary school with a diverse student population in a class with mixed-age students (Yannier et al., 2013).

Paper pre- and post-tests demonstrated large learning gains. On multiple-choice items asking students to predict which of two towers would fall first, 62% answered correctly at pretest, and 78% answered correctly at posttest ($t(11)=4.2$, $p<0.002$, $d=0.78$). On items asking students to explain why a tower fell first, 17% answered correctly at pretest, while 71% answered correctly at posttest ($t(11)=9$, $p<0.001$, $d=2.98$). Also, students were asked to build their own towers before and after interacting with the game. For all participants, the towers they built after playing the game were more stable than the ones they built before (Yannier et al., 2013).

Qualitative video data revealed that the children had high levels of engagement and excitement when the table shook and made the towers fall. They also had 'a-ha' moments after making wrong predictions and then seeing the explanation menu, which prompted reflection on what had happened. The children also seemed to collaborate productively: they discussed with and learned from each other. For example, while making a prediction they would explain to each other why they think one of the towers will fall first, making statements such as "Look! That one will fall first because it has a bigger top". Another example of collaboration and joint explanation development was when they were building towers together after interacting with the game. When one child first tried to put more blocks on one side of the tower his partner warned him saying: "No, don't put all the blocks on one side, that would make it unbalanced. We want it to be the same on each side" (Yannier et al., 2013).

I designed a new experiment to 1) provide a controlled test of whether physical experimentation in the context of EarthShake enhances learning, and 2) to probe hypotheses for why such learning benefits may occur. Qualitative data from the pilot suggested that physicality coupled with interactive feedback might play an important role, as it seemed to increase engagement and embodiment. Additionally, students' collaborations and discussions might have lead to learning (Yannier et al., 2013). To separate the factors of media-type and collaboration, I designed a 2x2 experiment: one factor contrasted EarthShake with a matched screen-based version of the game (mixed-reality vs. virtual), and a second factor contrasted collaborative and individual work (pair vs. solo).

If the benefits of physical observation stem from its enhancement of student collaboration, then we would only expect learning from EarthShake to be better than the virtual analog for the collaborative pairs. Alternatively, if physical observation fosters engagement and/or embodiment, which then yields greater learning, then we would expect better learning from EarthShake for both solo and pair groups. I include measures of engagement and embodiment to evaluate their potential roles in mediating learning.

4. EXPERIMENT 1 : MIXED-REALITY AND COLLABORATION

This experiment is designed to compare the effectiveness of mixed-reality and virtual conditions, which differ only in the medium of presentation: in the mixed-reality condition, students observe physical towers shaking and falling, while in the virtual condition students watch videos of the towers shaking and falling. Previous studies comparing virtual and tangible environments confounded the effects of touching and observing physical objects. This study isolates the effect of observation by ensuring that none of the students touch the towers while playing the game. All other important variables are tightly controlled (i.e., the role of the experimenter, the within-game and assessment questions, the game scenario, and the interactive feedback are kept the same). Only the medium of presentation is varied between conditions: virtual or mixed-reality (physical with interactive feedback).

As illustrated in Figure 1, this 2x2 experiment compared the mixed reality game EarthShake with an on-screen version of the same game (virtual) for solo vs. pair conditions. In the mixed-reality condition, the experimenter placed physical towers on the earthquake table. The game interface was projected onto a display screen directly behind the earthquake table. The gorilla character asked the students to predict which tower would fall first. Students made a prediction by clicking on one of the virtual towers, then observed the one of the physical towers fall. They then received feedback from the gorilla character, telling them if their prediction was right or wrong and prompting them to explain why this tower fell. Students selected explanations from a multiple-choice menu, as in the pilot study. In the virtual condition, instead of watching physical towers fall, students observed pre-recorded videos. To make the conditions as equivalent as possible, I videotaped the towers shaking on the earthquake table for each contrasting case in EarthShake. These videos were integrated into the game interface projected on the display screen. After watching the video, students received the same feedback and explanation prompts as in the mixed-reality condition. In both conditions, students used a mouse to interact with the interface. Additionally, since students in the target age group may not be fluent readers, all instructions, prompts, explanation items and feedback were read aloud with voice over by the gorilla in both conditions. The videos also included clear sound of the towers falling on the earthquake table. For the solo condition, the students interacted with the game on their own; in the pair condition, they discussed their answers with their partner before making a decision. For both the mixed-reality and virtual conditions, the experimenter sat next to the students but did not give any feedback.

The experiment had a between-subject design: participants were randomly assigned to a condition and interacted either with the mixed-reality or virtual game. Sixty-seven students (16 pairs, one group of 3, and 32 solo), ranging from kindergarten to 3rd grade, equally distributed among the different grades, participated in the experiment. Half of the participants were recruited through an email sent to their parents on a college campus mailing list. The rest of the studies were conducted in two different local elementary schools with a diverse student population. The participants recruited through the email list took part in the study in the lab, where as others participated in their schools. The pairs were either siblings or were selected by the teachers from the same class in the schools.

4.1 Procedure

Before playing, students independently completed a paper pretest to measure what they already knew about the stability and balance principles in the game. The experimenter helped with reading the questions and writing their answers in the paper tests for the students who had difficulty reading or writing. Next, students did a tower building task. They were asked to use a given set of blocks to build a tower that would stay up when the earthquake table shakes. Students were told to use a specific block as the base of the tower. Students in the pair conditions worked together to build one tower, while students in the solo conditions build their towers independently. Students then interacted with their assigned game, either EarthShake or the screen-only control. Each game included 10 contrasting cases (Figure 3). After interacting with their game, the students were given the same tower building task as before. This allowed us to measure the improvement in their towers after interacting with the game. After building the tower, they were given a matched paper posttest. Finally, the students took a survey which asked "How much did you like the game?". They choose one of: "I didn't like it at all", "I didn't like it", "It was OK", "I liked it", "I liked it very much". The first author also briefly interviewed the participants to see what they liked/disliked about the activity and if they had any suggestions. The same procedure was used for both the virtual and mixed-reality conditions.

4.2 Results of Experiment 1

I wanted to see the effects of mixed-reality and collaboration on learning and engagement. To accomplish this, I analyzed paper pre and posttests, tower pre and posttests, and the surveys that were given after the game.

I analyzed the results for the pre and posttests to identify any differences between conditions, media type (virtual vs. mixed-reality) and collaboration type (solo vs. pair). A 2-way ANOVA with overall pre-test score as the outcome variable confirmed no differences between the conditions at pretest (all F's < .79 and p's > .37). To test for learning benefits, I ran a 2-way ANCOVA with post-test score as the outcome variable and pre-test as the covariate. I found significant positive effects of the mixed-reality condition. The overall results indicated that the average scores on the full posttests (prediction and explanation items) was 64% for the mixed-reality condition and 48% for the virtual condition, $F(1,66)=23.3$, $p<0.0001$. The effect size of $d=0.78$ (Cohen's d) indicates a large effect. There was no effect of collaboration and no interaction effect of media-type and collaboration: the mixed-reality condition learned more than the virtual condition, both for pair and solo (Fig 5a). There was no significant difference in time on task between the four conditions.

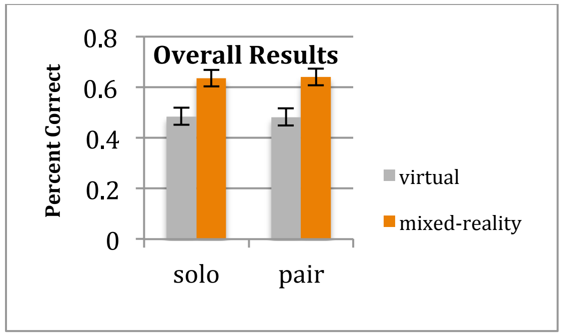

Figure 5a. Overall Post-test Learning Results

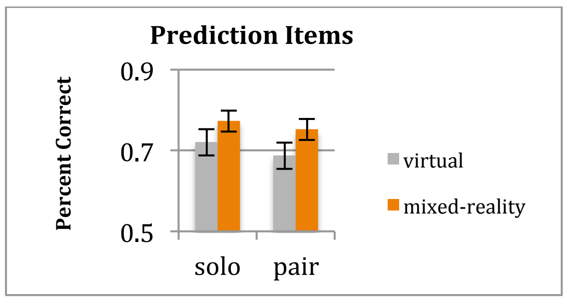

Figure 5b. Results for Post-test Prediction Items

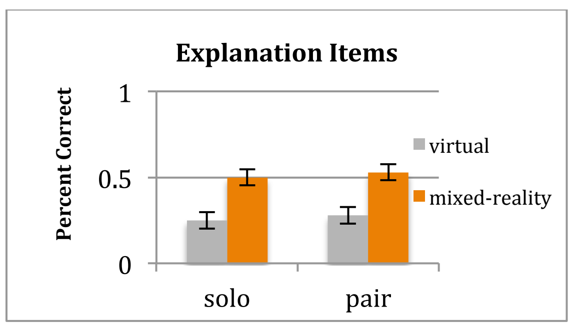

Figure 5c. Results for Post-test Explanation Items

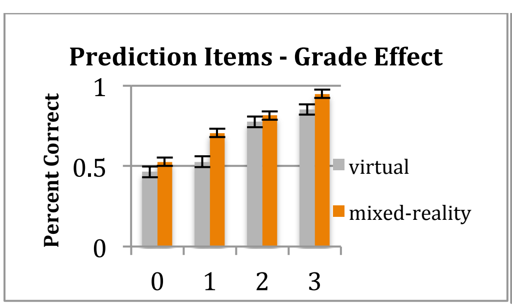

Figure 5d. Prediction Items Grade Effect

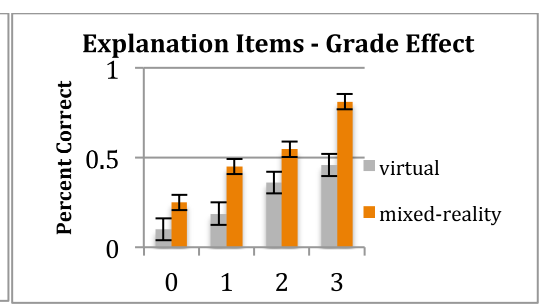

Figure 5e. Explanation Items Grade Effect

Considering only the prediction items, there was again a significant positive effect of the mixed-reality game. The average posttest score for the mixed-reality condition was 76%, while that of the virtual condition was 70% ($F(1,66)=3.1$, $p<0.0035$, $d=0.39$). Again, there was no effect of collaboration and no interaction effect of mixed-reality and collaboration: the mixed-reality condition improved prediction skills more than the virtual condition for both pair and solo groups (Fig 5b). The results were similar for the explanation items: A 2-way ANCOVA test showed that the mixed-reality condition learned significantly more than the virtual condition. The difference in this case was even higher (52% vs. 26% respectively, for posttest items, $F(1,66)=18.6$, $p<0.0001$, $d=0.87$). Note the large effect size (0.87) of this difference. Again there was no effect of collaboration and no interaction effect of mixed-reality and collaboration (Fig 5c).

I also analyzed the data by grade level. Across grades, higher grades performed better. Within each grade, students learned more in the mixed-reality condition, demonstrated both in the explanation and prediction items (Figures 5d and 5e). This finding that performance rises with grade level is evidence for the validity of my measures of learning. More interestingly, it provides an additional basis for estimating the size of condition effects in practical terms: namely, how much value the treatment condition adds relative to a year of schooling. The effect of grade is 9.5 points per year where as the effect of mixed-reality condition (over the virtual) is 9.4 points[1]. Thus, the treatment contributes 9.4 percentage points (relative to control), that is about equal to a year's worth of schooling, which contributes 9.5 points. This approach of using the whole year increases as a baseline for judging the size of a treatment has been increasingly used (Koedinger et al., 2010) and recommended (Lipsey et al., 2012).

To measure pre- to post-test changes on the tower building task, I scored each student's towers according to three principles: height, symmetry, and center of mass (I did not use the fourth principle, wide base, as all students were instructed to use the same base block). For each principle, students were given one point if their towers improved from pre- to post-test, -1 for the reverse, and 0 for no change. Comparing pre- and post- towers for the height principle, a shorter post-tower scores 1, a taller post-tower scores -1, and towers of the same height score 0. Likewise, post-towers with more symmetry and a lower center of mass score one for each of those principles. Adding the scores for each principle yielded the student's total score (Figure 6).

An ANOVA showed a significant effect of condition for the tower scores, in favor of mixed-reality ($F(1,66)=6.9$, $p=0.01$, $d=0.48$). There was no significant effect for group size (solo vs. pair) and no interaction effect of mixed-reality and group size. Thus, the children in the mixed-reality condition improved more on building stable towers than those in the virtual condition, for both the solo and pair conditions (Figure 7).

All three measures (the prediction items, the explanation items, and towers) showed a significant positive effect of mixed-reality conditions. What might explain this benefit? This thesis explores three likely mechanisms suggested by prior work: collaboration, engagement, and embodied cognition (Carini et al., 2006; Antle, 2012). Comparisons of the solo and pair conditions did not suggest any effect of collaboration. My quantitative and qualitative analyses, described below, provide evidence for embodied cognition but not engagement.

[1] This value of 9.5 points is the grade coefficient of a regression model with overall post-test as the dependent variable and interaction-type (virtual vs. tangible), grade, and pre-test as the independent variables such that, for example, a 2nd grader scores about 9.5 points higher than a 1st grader.

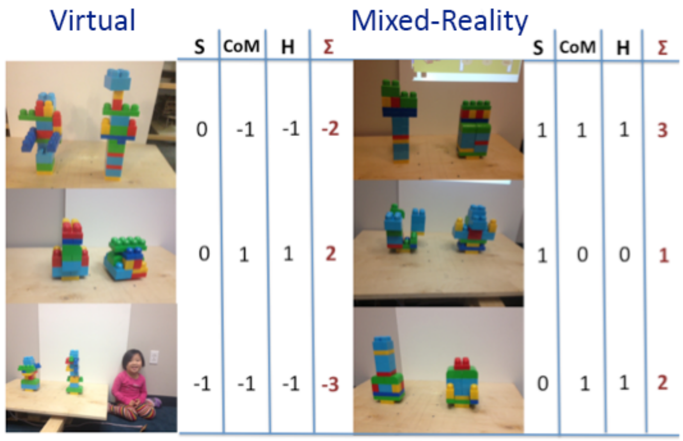

Figure 6. Coding scheme for Tower pre/post tests change.

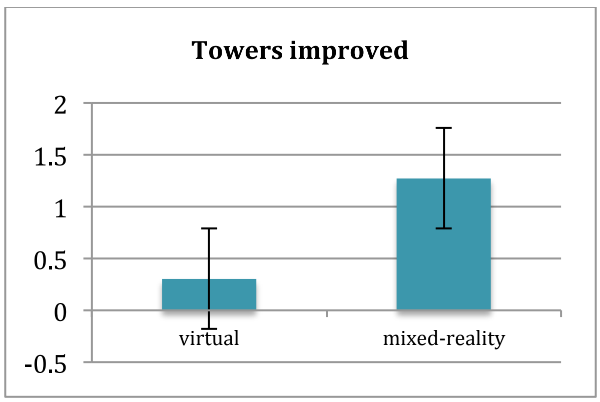

Figure 7. Scores on the tower building task (out of three). Positive scores indicate pre-to-post improvement in stability; a score of 0 indicates no change from pre to post.

4.3 Engagement and Enjoyment

Informal review of the video data suggested that children in the mixed-reality condition were highly engaged. They were especially excited when the live earthquake table confirmed their prediction of which tower would fall first. Some children even jumped up and down (see Figure 8). I did not see this level of engagement (e.g., jumping) in the virtual condition.

Figure 8. Engagement and excitement of children when they see that their prediction was right.

The formal survey (given after the post-test) provides another measure of enjoyment. Students were asked how much they liked the game, and responded with options on a likert scale 1-5 ("I didn't like it at all", "I didn't like it", "It was OK", "I liked it", "I liked it very much"). Students in the mixed-reality condition had higher mean ratings for enjoyment, and an ANOVA showed that this difference was significant ($F(1,66)=6.9$, $p=0.01$, $d=0.48$). There was no significant difference between the solo and pair groups for likability (See Figure 9. 1-5 scale was converted to proportion 0-1). However, the difference in enjoyment does not explain the difference in learning. Repeating the learning analysis only for students who gave the highest enjoyment rating still indicated a significant, favorable effect of mixed-reality ($p=0.001$).

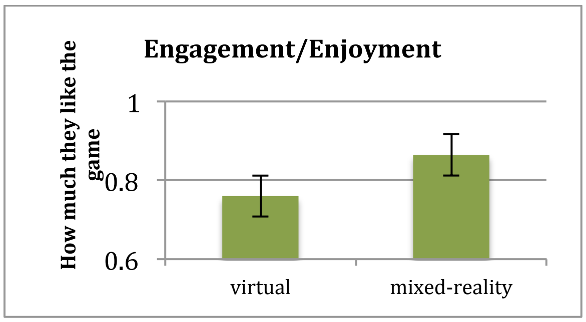

Figure 9. Results of the survey given to measure how much children enjoyed the game.

4.4 Gestures as Signs of Embodied Cognition

Based on Alibali's theory that gestures can be signs of people's mental visualizations and embodied language (Hostetter, Alibali, 2008), I used a measure of children's gestures as a proxy for embodied cognition. While analyzing the videos, I noticed that the children in the mixed-reality condition appeared to be using more gestures to explain their predictions. They were mimicking the tower structures and showing how the towers were structured with their hands. For example, while explaining his prediction of which tower would fall, one student said, "Because that one doesn't have a base, the base is just the same as the top." As he spoke, his gestures indicated the shape of the base. Another student explained "Because number one has a sturdier bottom," making a gesture suggestive of the length of the base (Figure 10). In the virtual condition, students mostly explained their predictions by pointing at the screen rather than using gestures that mirrored properties of the towers. An ANOVA analysis of the video data revealed that students in the mixed-reality condition used significantly more gestures than those in the virtual condition, when they were explaining their predictions ($p=0.001$, $d=0.72$). I counted only the gestures referring to the tower's structures, and did not count the pointing gestures in the analysis for any condition (Figure 11). For the statistical analysis, one participant from each condition was removed from the gesture analysis because their gesture counts were higher than five standard deviations above the mean.

Gestures invoking structure may indicate students' three-dimensional mental visualization. The finding that more of these gestures occurred in the mixed-reality condition suggests that seeing physical towers supports mental visualization better than seeing a video.

Figure 10. Children in the mixed-reality condition (above) used more shape-relevant gestures while explaining their predictions than those in virtual condition (below).

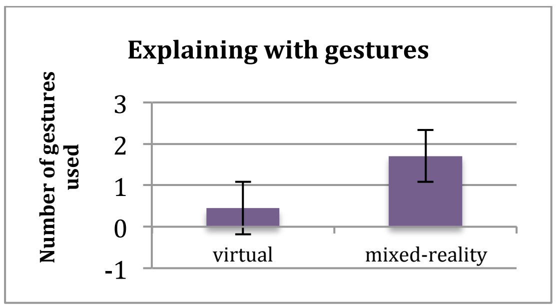

Figure 11. Average number of meaningful gestures used to explain predictions.

Looking at the data more closely, I saw that there was no significant correlation between gestures and learning. However, I realized that there was an interesting triangular structure in a scatter plot of gestures versus learning gains. There are many students who do not gesture, and some of these students do learn (i.e., show a positive increase from pre to post test in overall score on the paper test). See the "No Gesturing" column in Table 1. However, there are very few students who gesture but do not learn as shown in the "Gesturing" column in Table 1. This asymmetric pattern is statistically reliable (Fisher's exact test for asymmetry $p<0.05$). Thus, the data are consistent with the hypothesis that gestures are a sign of mental visualizations that enhance learning. If students do not gesture, they may nevertheless still be mentally visualizing, however if they do gesture, it is a sign of their mental visualizations that is associated with better learning.

	No Gesturing	**Gesturing**
Learning	0.34	0.23
No Learning	0.38	0.05

Table 1. Percentage of students gesturing (using meaningful gestures while explaining their predictions) and learning.

4.5 Qualitative Evidence

Qualitative anecdotes illustrate the students' enjoyment and engagement. Many children commented after the game that they liked the shaking table and the gorilla character. One expressed her enjoyment by saying: "It's so so much fun!" Another liked that the gorilla told him if he was right or wrong. Some commented that they liked guessing if the tower would fall or not. A mother of a participant said, spontaneously, that she would like to play the game at home, as a family and thought that it could help strengthen the family bond. Many children said that they would like to build their own towers and test them on the earthquake table, suggesting that hands-on activities may lead to even more engagement.

Further, while the pairs condition was designed to be collaborative, some students indicated that they would enjoy competing to build a tower that stayed up longest.

Most of the a-ha moments occurred after children made a wrong prediction and then recognized the relevance of one of the explanation options. For example, one child predicted that the left tower (Figure 4 – contrasting case 8) would fall first. Once the table shook, she saw that her prediction was wrong. When the multiple-choice explanation menu appeared on the display screen, she quickly selected her answer, exuberantly exclaiming "Ooooh because it has more weight on top than bottom!" I suspect that observing the physical outcomes rather than the video leads children to take evidence against their prediction more seriously and thus more actively engage in trying to find an explanation. One child in each condition commented that they would prefer seeing the towers fall in real life rather than having a video or the computer say what happens.

It was also interesting to see that the children interacted very naturally with the interface. They just assumed that the gorilla could see the physical towers and did not even realize there was a camera in the setup. This was what we had wanted, as good technology should be transparent.

One of the moms indicated that she believed the game might help family cohesion. She suggested that it might be even better if the game involved both collaboration and competition, such as having two teams consisting of the mom and child vs. dad and child competing with each other. I also observed that children enjoyed competing in some cases. For example after the game was over, one of the pairs wanted to continue playing. They wanted to build their own individual towers (rather than one collaborative tower together as given in the task) and test whose tower would stay up longer on the table.

5. EXPERIMENT 2 : MIXED-REALITY AND PHYSICAL CONTROL

Experiment 1 showed greater learning gains when children observed physical towers rather than watching videos of the same. This result suggests learning benefits for young children from physical observation, even when students do not touch the objects. Experiment 2 explores if adding a simple and scalable physical control (such as shaking a tablet) could further increase learning by increasing enjoyment or if physical observation in the context of a mixed-reality environment is more critical to learning.

Experiment 2 replicated the mixed-reality vs. screen-only comparison from Experiment 1, and crossed each condition with the presence or absence of a simple physical control (Yannier et al., 2015). Since Experiment 1 found no differences for learning or enjoyment between the solo and pair conditions, all participants interacted with the game in pairs in this follow up experiment. Experiment 2 used the same tests and surveys as Experiment 1 to measure enjoyment and learning gains. The physical control in the mixed-reality game consisted of a physical switch that the children pressed to shake the table. For the screen-based version, the game was implemented on a tablet, which children physically shook to shake the virtual table. In pilot tests, children seemed excited about pressing the physical switch to shake the table, suggesting that a physical control would lead to greater enjoyment.

5.1 Procedure

I developed the technologies that would be used in the four experimental conditions: 1) mixed-reality version of EarthShake with mouse control; 2) mixed-reality version of EarthShake with physical control (pressing a physical button as input); 3) Screen-only laptop version of EarthShake with mouse control; 4) Screen-only tablet version of EarthShake with physical control (shaking the tablet as input). In each condition, students played in pairs. I discuss each in more detail below.

5.1.1) Mixed-reality version of EarthShake with Mouse Control

This condition was equivalent to the mixed-reality & pair condition in Experiment 1. In this condition, children indicated their prediction of which tower would fall by clicking one of the choices on the projected screen. Then, the children clicked a 'shake' button, also on the projected screen. After the children made this selection, the experimenter used a physical control to shake the earthquake table.

5.1.2) Mixed-reality version of EarthShake with Physical Control

This condition is identical to the mixed-reality version of EarthShake with mouse control, except that the children were given the physical control (a physical switch connected wirelessly to the earthquake table) to shake the table (See Figure 12).

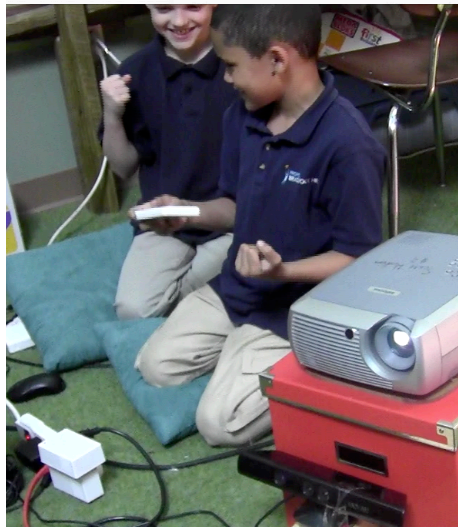

Figure 12. Students using a physical switch to shake the table while interacting with the mixed-reality version of EarthShake with physical-control.

Each child in the pair took turns holding the physical switch, which shook the table, and using the mouse, which controlled the prediction and explanation selections. To ensure that the child only shook the table after a prediction was selected, the experimenter wirelessly disabled the child's switch until the appropriate time.

5.1.3) Screen-only version of EarthShake with mouse control

This condition was the same as the virtual & pair condition in Experiment 1. The participants used a mouse to control the game on the screen. They were asked to take turns using the mouse.

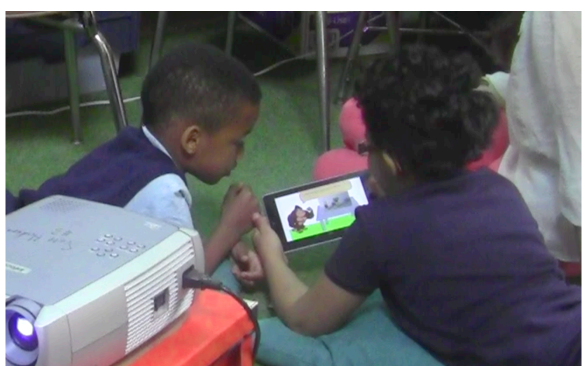

Figure 13. Virtual & physical-control condition where children shake the tablet to shake the table on the screen.

5.1.4) Screen-only tablet version of EarthShake with physical control

In this condition, children used a tablet version of EarthShake. This implementation included the same game interface, gorilla character, scenario, and button controls as the mixed-reality and the laptop versions. Like the laptop version, a video of the towers was integrated into the game interface. Unlike the laptop version, the tablet version included a physical control: children shook the tablet with their hands to activate the video of the towers falling (Figure 13). In this condition, the partners were asked to sit on the floor next to each other in a way that would allow both of them to see the screen of the tablet. They took turns shaking the tablet and clicking on the selection choices.

The experiment had a between-subject design, with each pair of students randomly assigned to a condition. Ninety-two 6-8 year old children, grades K to 2 participated in the study (43 pairs and two groups of 3). Children were recruited from two different schools with a high percentage of students from low-income communities. The pairs to take part in the experiment were randomly selected by the teachers. The same procedure as in Experiment 1 was used. The same measures were used as in Experiment 1.

5.2 Results of Experiment 2

Paper pre and posttests and tower pre and posttests were analyzed to measure the learning gains from the experiment and investigate the effects of observing physical phenomena and physical control on learning. Surveys were analyzed as a measure for enjoyment.

A 2-way ANOVA analysis with overall pre-test score as the outcome variable was performed revealing no differences between the conditions at pretest (F's < .46 and p's > 0.50). To investigate learning benefits, a 2-way ANCOVA was conducted with between-participant factors of control-type (mouse-control or physical control) and media-type (mixed-reality or screen-only), with pre-test score as a covariate and post-test as the outcome variable. The overall results (including both the prediction and explanation items) indicated that there was a significant effect of media-type ($F(1,91)=8.2$, $p<0.01$, $d=0.37$), with benefits for mixed reality. The average score on the posttests (both the prediction and explanation items) was 45% across the mixed-reality conditions and 39% across the virtual conditions. The overall improvement from pre to post was 11.3 % in the mixed-reality conditions and 2.4 % in the virtual conditions, revealing that the mixed-reality game improved learning by 4.8 times compared to the screen-only alternatives. No significant effect was found for control type and there were no significant interaction effects. Thus, mixed reality led to more learning than screen only, for both the mouse-control and physical-control conditions (Figure 14). This result indicates that, for young children, physical observation can improve learning, while simple physical control is unlikely to.

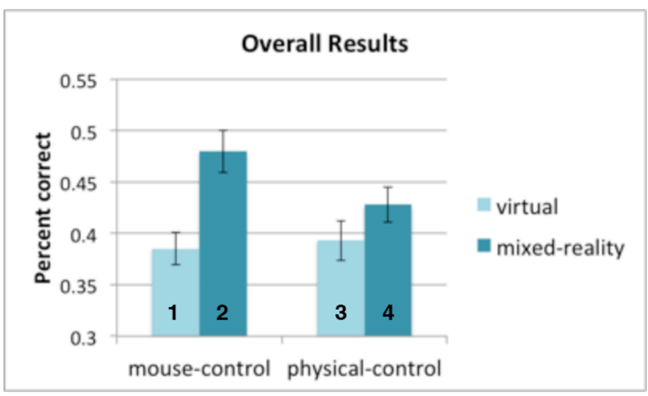

Figure 14. Overall post-test results.

Considering the conditions separately, we can see that the mouse-control mixed-reality condition (#2 in Figure 14) is significantly better than a typical virtual (#1) ($p<0.05$), while the virtual, physical-control condition (#3) is not. Thus, facilitating physical observation was more powerful than facilitating physical control through shaking the tablet for learning.

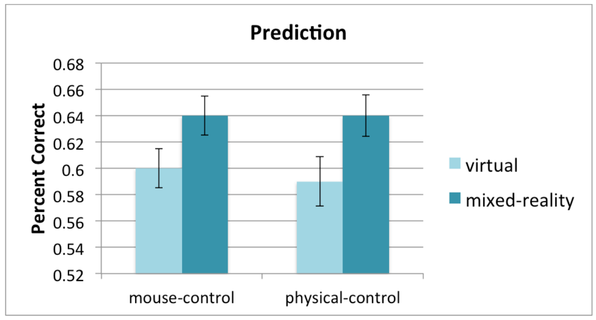

Figure 15. Post-test scores for Prediction Items

The main effect of media type, in favor of mixed reality, held for both the prediction and explanation items separately. The analysis for the overall scores was repeated for the pre- and post-test prediction items. Collapsing the conditions by media type, the improvement from pre to post for the prediction items was 7% for mixed-reality and was 1% for virtual ($F(1,91)=4.2$, $p<0.05$, $d=0.41$). The average posttest scores for the mixed-reality and virtual conditions were 64% and 60%, respectively (Figure 15). There was no significant effect of control-type and no significant interactions.

Likewise, for the explanation items, a 2-way ANCOVA showed significant differences in learning by media type, with the mixed-reality condition scoring higher at post-test than the virtual condition (Figure 16; 27% vs. 18% for posttest items, $F(1,91)=4.7$, $p<0.05$, $d=0.44$). The pre-to-post improvements in explanation items for the mixed-reality and virtual conditions were 15.5% and 3.7%, respectively. As with the overall scores and prediction scores, there was no significant effect of control type and no significant interactions. While the interaction between control type and media type is not significant, we do observe a trend: for students with the mixed-reality game, the, mouse-control condition was slightly better than the physical-control condition. One explanation for this trend could be that pressing the physical switch was so exciting for the children (supported by the data in the enjoyment section below) that they did not pay full attention to the explanations provided in the game.

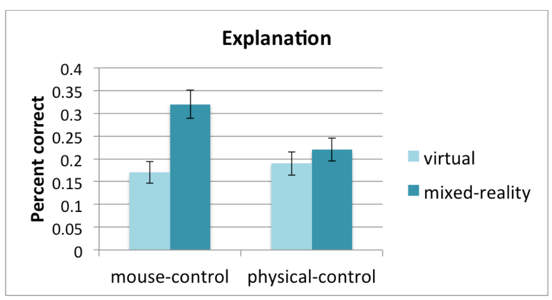

Figure 16. Post-test scores for Explanation Items

The pre and post towers were scored with the same coding scheme that was used in Experiment 1 (pre-to-post improvement scores are shown in Figure 17).

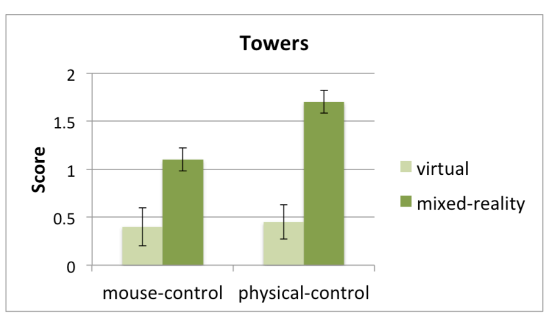

Figure 17. Tower scores

A 2 way ANCOVA showed that there was a significant effect of media-type for the tower scores, in favor of mixed reality (F(1,91)=6.9, p=0.01, d=0.64). There was no significant effect for control-type and no interaction effect of media-type and control-type. Students in the mixed-reality conditions improved their towers more than students in the virtual conditions, for both the mouse and physical control. This result is interesting as it shows that the benefits of physical observation over video-watching transfer to a constructive problem solving task involving physical interaction with the blocks as well.

5.3 Engagement and Enjoyment

Enjoyment was measured with the same survey as in Experiment 1 (Figure 18). An ANOVA on the survey results showed a significant difference in enjoyment by media type, with the mixed-reality condition indicating more enjoyment (F(1,92)=6.7, p=0.01, d=0.55). There was no significant effect of control-type for enjoyment. There was also no interaction effect of media-type and control-type. Though the interaction was not significant, we do observe a trend among students in the mixed-reality conditions: the physical-control students indicated greater enjoyment than the mouse-control students. Analyzing only those who interacted with the mixed-reality game (either with mouse-control or physical-control), I saw that there was a marginal effect of control-type (p=0.08, n=45): the physical switch in the mixed-reality game was increasing enjoyment slightly. On the other hand, scrutinizing only those interacting with the virtual conditions, control-type does not appear to have any effect.

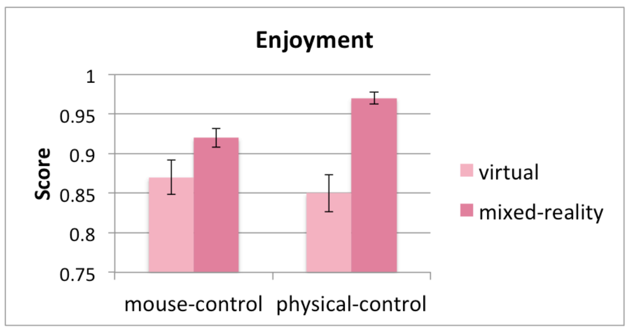

Figure 18. Enjoyment scores based on the survey

5.4 Gestures and Signs of Embodied Cognition

I examined the gestures students made while explaining their predictions and the results (why one of the towers fell first). To measure indications of embodied cognition and mental simulations, I coded those gestures with the same scheme as in Experiment 1. An ANOVA shows that the students in the mixed-reality condition were using significantly more meaningful gestures than those in the virtual condition, while explaining their predictions (F(1,92)=11.55, p=0.001, d=0.72) (Figure 19). This result is consistent with the gesture results from the first experiment.

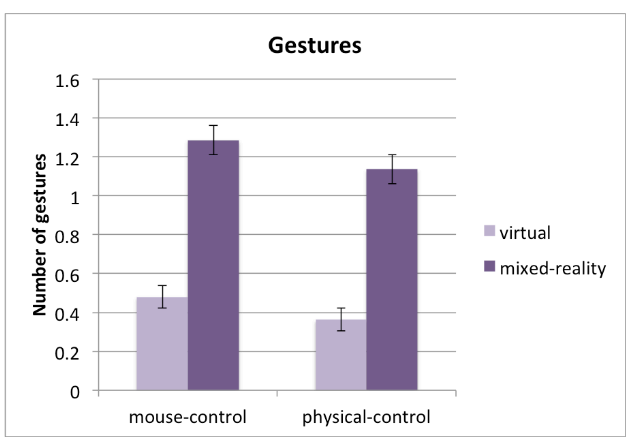

Figure 19. Meaningful gestures children used during their explanations.

Furthermore, there is a significant correlation between these meaningful gestures and overall learning gains ($R=0.21$, $p<0.05$). It is hypothesized that children's spontaneous gestures reflect their mental simulations and processes (Hostetter, Alibali, 2008). The significant correlation in my results is consistent with the hypothesis that mental simulations lead to more gestures and enhanced learning. Taken with the significant difference in gesture frequency between the mixed-reality and virtual conditions, it is likely that greater learning for students who observed the physical towers was due to better mental visualization.

5.4 Qualitative Evidence

The qualitative evidence in this experiment was similar to that in Experiment 1. Again, children seemed to be very much engaged during the game. One of the children in the mixed-reality condition asked if she could trade some of her toys to get EarthShake. Another student asked if she could steal the experimenter's computer to set it up at home. Others made remarks saying that they never thought something they do at school could be so much fun and that they wished all their science classes were fun like this. Another student cited that she thought this was like the next version of smart boards.

Furthermore, there were some children who imagined shapes from the physical towers and started making up stories. After seeing one of the towers fall, one student started laughing, stood up and said: "It's like a giraffe. It falls after the earthquake… and that's the tree" pretending to be a tree with her arms wide open. Thus, seeing real towers may be triggering children's imagination, facilitating embodiment of the stories, helping them make connections with objects they are already familiar with and as a result assisting their learning.

6. DISCUSSION OF EXPERIMENT 1 AND 2

EarthShake uses an affordable camera and a projector to combine the advantages of the physical and virtual worlds. In doing so, it presents a new kind of mixed-reality learning that incorporates prediction, physical observation, explanation, and personalized immediate feedback. Experiment 1 revealed significant differences in learning between the virtual and mixed-reality conditions, as measured by greater pre to posttest gains in predictions, explanations, and constructed towers. Experiment 1 found no significant differences between the solo and pair conditions. These results demonstrated the benefits of observing physical phenomena over watching a video of the same event. From a theoretical perspective, this finding begins to tease apart the effects of observing versus manipulating physical objects – factors that have not been controlled for in previous work. Experiment 2 replicated this intriguing result and also showed that a simple hands-on control, such as shaking a tablet or pressing a switch, does not have a significant effect on learning or enjoyment.

Furthermore, my results revealed that the learning benefits transfer from prediction to a construction task as well. The towers of the mixed-reality conditions improved significantly more than those in the virtual conditions. Thus, students were not only better learning physics principles of balance, but they could also better apply them in a constructive problem solving task involving hands-on manipulation.

As far as we know, this thesis presents *the first* randomized controlled experiments showing that physical observation in the context of an interactive game can improve enjoyment and learning for children above and beyond an equivalent screen-based tablet or computer game. It may be that touching or manipulating the towers has a further learning benefit and that is a question for further research. Nevertheless, these results show that observing physical towers accompanied by interactive feedback, in of itself, has a strong effect on enhancing science learning.

Why did the mixed-reality game lead to better learning? I explored three theoretical explanations for why observing the physical phenomenon may produce more science learning: that physicality is inherently more engaging, that it facilitates embodied cognition, and that it enhances collaboration. My data does not support the collaboration theory, as the pairs in Experiment 1 did not learn more than the students playing solo.

Another possible explanation is that students learned more from the mixed-reality condition because of their increased enjoyment and engagement. Students in the mixed-reality condition both qualitatively showed more engagement and rated their engagement higher on the quantitative survey. However, while the data supports the theory that physical objects are more engaging than virtual ones, increased engagement does not seem sufficient to explain the large learning differences I observed. To check for such an effect, I analyzed the subset of participants who each gave their game the maximum likability rating (14 in the mixed-reality condition, 10 in the virtual), and still found a significant effect of media-type on learning ($p=0.001$). This result was replicated for both experiments.

The gesture data provides some support for the explanation that physicality supports embodied cognition and triggers affordance for action, which as a result helps children perceive, mentally visualize, and ultimately remember concepts better. Children in the mixed-reality conditions more often explained their predictions using meaningful gestures to show 3D motion than children in the virtual condition. This finding suggests that those students had mentally visualized the objects, which may have helped them register and remember the explanations for why each tower fell. This result is in line with prior work that suggests (1) when children learn abstract concepts, they utilize mental simulations based on concrete motor-perceptual experiences (Antle, 2012); and (2) the spontaneous gestures children spontaneously produce when explaining a task are a sign of their mental visualizations and predict how much they will learn from that task (Cook et al., 2008).

In sum, the current evidence is most supportive of the theory that physicality supports mental visualization and enhances retrieval and reasoning through embodied cognition. The evidence is perhaps least consistent with the idea that the results are merely a consequence of increased enjoyment.

7. DESIGN IMPLICATIONS FROM EXPERIMENT 1 AND 2

The design iterations I made on EarthShake revealed the importance of having a well-planned sequence of guided-discovery activities (including a predict-observe-explain structure and contrasting cases) in conjunction with a self-explanation menu and interactive feedback that scaffolds students to construct their own explanations and understanding of early physics principles without being told directly. In the pilot studies where I used the earthquake table on its own, without the projected game, it appeared children were having less success in learning the physics principles. In contrast, when they saw the self-explanation menu while also seeing what happened in the physical world in the foreground, they were able to recognize the principle that was causing the phenomena (such as having more weight on top than bottom) even if they had not predicted it beforehand. Thus, I believe that the self-explanation menu synchronized with the physical world was a critical component of the game and facilitated learning.

I realized that kids liked the hands-on activities and wanted to have more building integrated into the game. They mentioned that they enjoyed building their own towers and testing them on the earthquake table. One child explicitly indicated that he would like it better if the game had more building. Thus, incorporating more hands-on activities in the central game mechanic (and addressing the associated technical challenges) may yield further benefits.

Some of the children complained that there was too much voice over, especially when the gorilla read all the answers in the menu one by one (which was a design choice I made so that they would hear all the answers without skipping through them). One of the children expressed this complaint by saying: “I don’t want the gorilla to speak so much!”

I observed that some of the children had a hard time using the mouse. In some cases the single mouse created a barrier against collaboration since some children had trouble sharing and tried to grab the mouse from their partner. I believe it might improve the interaction if a more tangible approach was taken for the selection of menu items instead of the mouse, that is, by allowing students to select items by pointing at the screen or by physical manipulation.

It is important to emphasize that there are important elements of technology in the mixed-reality system that can provide benefits above and beyond having children simply play with blocks on an earthquake table. The Kinect camera and the specialized computer vision algorithm in the setup allow the system to provide task guidance (asking students to make a prediction, observe the results and reflect on what happened) and to give interactive feedback. In particular, the vision algorithm detects when an experiment is over (when one of the blocks has fallen), determines whether the child’s prediction was accurate, and gives feedback to the child that they can then use to make sense of the outcome. The gorilla character encourages self-explanation, asking the students to make a prediction, giving them feedback if their prediction was right or wrong and asking them to reflect on why, all synchronized with the real world via depth camera sensing. The explanation menu that appears in the projected game also scaffolds children in reasoning about the physical properties that cause stability. I observed several engaged 'a-ha' moments for students in the mixed-reality condition. For example, after watching the table shake, a student realized her original prediction was incorrect. Upon seeing the self-explanation prompts, she yelled "Oooh because it has more weight on top than bottom!" I did not observe any a-ha moments in the virtual condition. Thus, the explanations in the projected game scaffold students to understand the underlying principles.

Using this system (utilizing depth camera sensing to provide synchronized personalized feedback), more interactive feedback can also be added (after letting the students discover principles on their own), which may help explain and visualize the physics principles that are important (e.g. explaining with visualizations and animations why the asymmetrical tower fell etc.).

The feedback provided in the game is critical for three reasons: 1) There is much evidence that children learn better with feedback. It has been shown that guided feedback and self-explanation can improve learning (Aleven, Koedinger, 2000). There is also research showing that without scaffolding and support, people often miss the point of the learning activity (Puchner et al., 2001), and 2) in this particular domain of science, the phenomenon of "confirmation bias" (Nickerson, 1998) suggests that children are likely to see their predictions as confirmed even when they are not, so explicit indication otherwise can reduce this tendency. Thus my system utilizing depth camera sensing to provide personalized immediate feedback on top of the real world, allows children to discover new principles with some support and scaffolding.

8. PROPOSED WORK

My previous research suggests that having a mixed-reality environment bridging physical and virtual environments produces enhanced learning and enjoyment. However it is not clear, from previous literature, what features of a mixed-reality system are most important for maximizing learning and enjoyment and, in fact, many researchers advocate for features that are at odds with the current design of EarthShake. As mentioned above, most tangible interfaces and mixed-reality environments to date have been purely exploratory and there aren't enough experiments that investigate how they should be designed to enhance learning. There is research, which suggests that tangible interfaces are best suited for exploratory activity and learning through a process of discovery (Marshall, 2007). An example of a tangible interface that affords exploratory activity is Underkoffler and Ishii's (Underkoffler, Ishii, 1998) Illuminating Light, which is designed to enable the rapid prototyping of optical layouts. It is often claimed that tangible interfaces are particularly for exploratory learning, as interaction with tangible systems is found to be more natural or intuitive to students than other types of interface, affording a particularly suitable environment for rapidly experimenting. However, little comparative work has been carried out, and it remains unclear which elements of tangible interface designs are critical in supporting learning activities (Marshall, 2007).

With my proposed work, I aim to investigate the effect of exploration learning and guided discovery activities (using a predict/observe/explain/feedback structure) on learning and enjoyment in tangible interfaces and mixed-reality environments. Below I give some background about the literature and learning theories on guided discovery and exploration.

Learning Theories behind Guided Discovery and Exploration:

Pure discovery and exploration learning is based on (some interpretations of) constructivist learning theory, where open-ended discovery learning and hands-on exploration is believed to lead to better learning (Papert, 1980). Based on inquiry, discovery learning expects the learner to construct his/her own learning agenda. By experimenting and wrestling with results, the learner constructs new learning based on experience. Some have argued that pure discovery leads to better learning compared to guided instruction in Tangible User Interfaces (Schneider et al., 2015). Kirschner, Sweller and Clark criticize that discovery learning defies the cognitive load theory – the learners cannot retain the amount of information needed to process the content (Kirschner et al., 2006). Other critics have also pointed out the possibility of undetected misconceptions, student frustration and lack of worked examples (Tuovinen, Sweller, 1999).

In contrast to open-ended discovery learning, guided discovery combines the action of discovery learning with the aid of scaffolding to produce a rich, blended learning experience that recognizes the boundaries of cognitive load while encouraging the passion of student exploration. Students engage in active learning fueled by inquiry, where the missteps by the student are caught and redirected to ensure correct information is placed on the foundation of existing student knowledge. Providing the learner with plenty of opportunities for deliberate practice and reflection, guided discovery provides supporting information to search the content to satisfy the hungry mind and satisfy curiosity (Harvel, 2010). Kozulin, Gindis, Agevey and Miller call guided discovery a middle ground between linear didactic teaching and open-ended discovery learning (Kozulin et al., 2003). Alfieri et al.'s findings suggest that unassisted discovery does not benefit learners, whereas feedback, worked examples, scaffolding and elicited explanations do (Alfieri et al., 2011).

With this work, I aim to discover if guided discovery, exploration or a combination of both leads to better learning in mixed-reality environments.

8.1 New Design of EarthShake

In order to test the effect of exploration and guided discovery in a mixed-reality environment, I have decided to add more exploratory learning activities to EarthShake. My qualitative analysis from the experiments suggest that it may have benefits to have more hands-on activities in the game where children get to manipulate the physical objects and do more hands-on exploration (supported by Montessori's theory that young children are highly attracted to sensory development apparatus and that they use physical materials spontaneously, independently, and repeatedly with deep concentration (Montessori, 1964). My gesture analysis, showed that children have mental visualizations when they experiment/observe physical objects, leading to better learning. So it may be argued that having more hands-on manipulation and building activities with physical objects may lead to even more mental visualizations and thus better learning. From the surveys, I also saw that kids enjoyed building their own towers and testing them on the earthquake table. Thus based on the feedback and literature about hands-on learning, I decided to add more hands-on activities into the game. The new design of the game consists of two different modes: 1) Guided Discovery with Predict/Observe/Explain/Feedback cycle, 2) Hands-on Exploration and Problem Solving. It starts with a main menu, where users can choose either to play the game, or build their own tower. If they choose "play the game" option then they go directly to the guided discovery activity where they are asked to place two towers on the table and make a prediction about which one will fall first. On the other hand, if they choose "make a tower" option, then they go to the hands-on exploration activity where they can build their own tower and test how long it will stay when the earthquake table shakes. They can go back and forth between different modes, and are asked if they want to go to the other mode after interacting with the game a few times.

8.1.1 Guided Discovery with Predict/Observe/Explain/Feedback

In this mode, first the users are asked to place the towers (contrasting cases) shown on the screen on the table based on their color and shape configurations. The users place the towers on the table. If they place the wrong tower, the gorilla character on the screen tells them to place the correct tower kindly. If they place the correct tower, they are told "Good job". Once they place the correct towers, the gorilla asks them which of the towers will fall first when the table shakes. They make a prediction and choose one of the towers. Then the gorilla character shakes the table and tells them if their prediction was correct or not, also asking them to explain why they think

8.1.2 Hands-on Exploration and Problem Solving

In this mode, first the users are asked to build a tower that will withstand the earthquake when the gorilla shakes the table. When they make a tower and place it on the table, the virtual visualization of the tower is shown on the screen. The users can then shake the table by clicking the "shake" button. When the tower falls, the Kinect camera and our vision algorithm detects it and then gives feedback to the user, displaying the time it took for the tower to fall. If it stays up then the gorilla congratulates them saying "Good job!". They can also be given different challenges such as making a tower that is taller than a certain height.

8.2 Pilot at the Children's Museum of Pittsburgh

I have created the new design of EarthShake and then had two play testing sessions at the Children's Museum of Pittsburgh to get feedback from families and children and iterate on the design according to the feedback. In the first pilot, I got a lot of interesting feedback from the families. It seemed like the parents were very much involved in helping the kids. At some points, they even started playing the game themselves. The kids also seemed to be very much engaged. A child who had a birthday party that day, played with the game early in the morning before anyone else came, then she brought her whole birthday party to play later on.

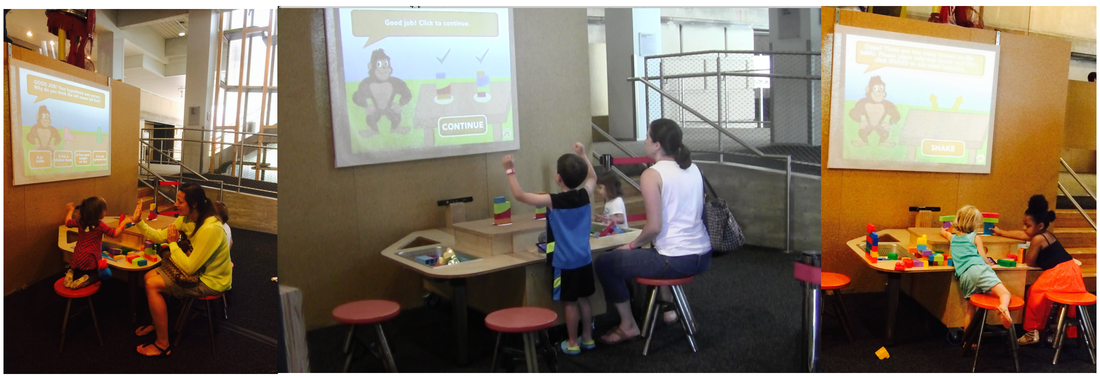

Figure 20. Pilot testing at the Children's Museum of Pittsburgh

8.2.1 Qualitative Data and Findings from Museum Pilot

Usability/design issues:

One issue I found was that the younger kids had a hard time using the mouse. Therefore I decided to use a tablet as an input device instead of the mouse. Also, our computer vision algorithm that was dependent on the colors of the blocks seemed to be effected by the lighting conditions in the room. Therefore, I decided to change the algorithm so it would use depth and shape information instead of color. I came up with a new algorithm that calculates the physics inertia of towers using their shape to identify which tower has been put up. Also, I saw that the prebuilt towers that I had created by sticking lego blocks together were not durable. Children thought that they could be taken apart when they saw the lego blocks (glued together to create prebuilt towers) and tried to separate them. However, I realized that the colors of the towers were actually very helpful for kids and parents to distinguish between the towers and use as cues to discuss/decide which tower to put up. Thus I decided to create prebuilt towers made of wood blocks that would be secured tightly to each other and painted to create a visual cue.

Findings:

I had some interesting observations from the pilots. I observed that parents and children seemed to discuss and engage in productive dialogue, discussing the reasons why the towers fall, more in the guided-discovery condition. When the game prompted them to explain why they think one of the towers fell first, the parents started asking children why they think the tower fell and helped them understand the underlying principles. For example, one of the moms put the towers side by side, asking her son what the difference is between the two, if one has more weight on top than the other. Similarly, many other parents got involved and tried to guide the kids to understand why.

I also observed that children and families seemed to strategize more while building their towers when they did it after the guided discovery activity as opposed to doing it as a first activity when they first came to the exhibit. If they chose to build a tower before interacting with the guided-discovery activity, they tended to do more random things that are not aligned with the goals of the game, e.g. trying to make a tall tower that will fall quickly.

I also saw that people tended to spend more time with the guided-discovery activity than building a tower. However this also depended on which activity the person before them was interacting with. If they saw somebody playing the game, then they started playing the game and continued with it. On the other hand if they saw somebody building a tower, they tended to start building a tower, too.

In the section below I describe the two parts of my proposed work, 1) an experiment that addresses my main research questions 2) design of the mixed-reality platform NoRILLA and different games that can be extended to different content areas in education in museum and school settings.

8.3 Proposed Experiment: Guided Discovery vs. Exploration/Problem Solving

8.3.1 Experimental Design

Although there are many mixed-reality environments and tangible interfaces for learning, it's not clear whether these interfaces are more beneficial for children and, if so, what features are critical to create a mixed-reality environment that is optimized for learning and enjoyment. Most of the current tangible interfaces are purely exploratory or based on open-ended problem solving, but do not have a strong pedagogical guidance provided by intelligent feedback for the users. In my proposed work, I want to explore whether having an interactive guided discovery layer, made possible through depth-camera sensing and AI perception of the real world, can improve children's learning and enjoyment compared to pure exploration. Further, might a smart combination of guided discovery and exploration further help improve learning and enjoyment?

These questions lead naturally to three experimental conditions as shown in Table 2. For the Guided Discovery condition (GD), children will be given about 10 contrasting cases, for which they will be asked to make a prediction, observe and explain the results with interactive feedback as explained in section 8.1.1. For the Exploration condition (E), children will be asked to build towers that will stay up as long as possible when the table shakes. As explained in section 8.1.2, this condition will involve limited interactive feedback that informs children if their tower stayed up or not and for how long it stayed up. They will be asked to build multiple towers for a time period that is equivalent to the time required in the Guided Discovery condition. For the Combined condition (GDE), children will be given 4 contrasting cases (targeting different principles), followed by an exploration task (building their own tower), and then 4 more contrasting cases followed by an exploration task. Pilot testing will be used to inform the best combination and to be sure that learning time across conditions is well matched.

Conditions:

Condition	**Explanation**
Condition 1: GD	Guided Discovery (8.1.1)
Condition 2: E	Exploration/Problem Solving (8.1.2)
Condition 3: GDE	Combined Guided Discovery + Exploration (8.1.1 & 8.1.2)

Table 2 Different conditions that children will be given during thee experiment.

Research Questions:

The experiment helps us answer the main research questions:
How to make a mixed-reality environment that is optimized for learning and enjoyment:
1) Does guided discovery using predict-observe-explain-feedback structure help kids learn better compared to exploration/problem solving in a mixed-reality environment or is a combination of exploration/problem-solving and guided discovery better for increasing fundamental concept learning and enjoyment? (Condition 1(GD) vs. Condition 2(E) vs. Condition 3(GDE))
2) Does guided discovery, exploration or a combination of the two transfers better to a hands-on tower building task? (Condition 1(GD) vs. Condition 2(E) vs. Condition 3(GDE))

8.3.2 Participants

I will conduct the study with 4-8 year old children in a lab or school setting with at least 20 children in each of the three conditions. Children will be recruited from a school or via email list. Children will interact with the game in pairs.

8.3.3 Procedure

Before interacting with the game, students will first be given a tower task. They will be asked to use a given set of blocks to build a tower that would stay up when the earthquake table shakes, using a specific block as the base of the tower. Then they will complete a paper pretest to measure what they already knew about the stability and balance principles in the game. Next, children will be given a tower building task. Students will then interact with their assigned game, either guided-discovery (GD), exploration (E), or the combined guided-discovery and exploration (GDE) condition. The condition content will be designed through piloting to involve the same time and some maximum time count-off will be employed in the case some children take much longer than expected. After interacting with their game, students will be given a matched paper posttest. After the paper posttest, the students will be given the same tower building task as before game play. As in prior studies the contrast between the pre- and post- tower building assessment will be used to measure student improvement in tower building, and in particular, how well they incorporate the principles of balance. Finally, the students will be asked to fill out a survey to see how much they enjoyed the game.

8.3.4 Measures

The paper pre and posttests are prepared based on the NRC Framework & Asset Science Curriculum (Quinn et al., 2012). The tests consist of two types of items: prediction items and explanation items. For the prediction items, the students will be given a picture of a table with two towers and will be asked to predict which will fall when the table shakes. In the explanation items, they will be asked to explain why they chose their answer (Figure 21). Also, children will be given a survey at the end of the game to measure their enjoyment. The survey will ask “How much did you like the game?” They can choose one of: “I didn’t like it at all”, “I didn’t like it”, “It was OK”, “I liked it”, “I liked it very much”.

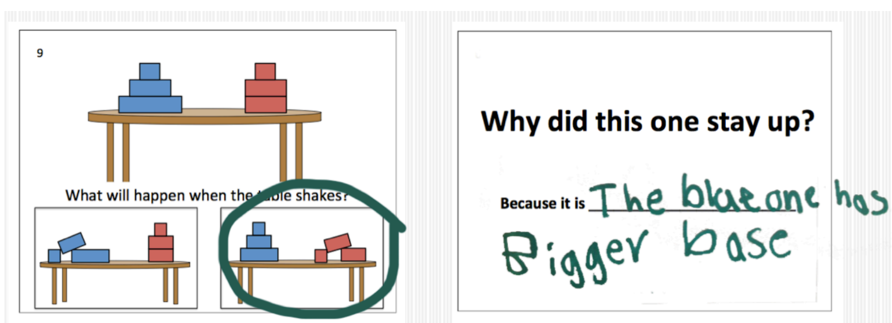

Figure 21. Prediction (left) and explanation (right) items used in the paper pre/posttests.

8.3.5 Hypotheses and Data Analyses

Alternative hypothesis for how learning and enjoyment can be maximized can be inferred from existing theories and learning-science based recommendations. Table 3 outlines how these alternatives yield different predictions for assessments of different outcomes.

Perhaps one of the most straightforward claims about learning is that you learn what you practice. An associated strong instructional design recommendation is that instruction should be aligned with the learning goals as made concrete and observable through targeted assessment tasks (Carver et al., 2010). If you want students to learn to do tasks of a certain type (say type A), then give students practice on tasks of that type (type A) and do not give them practice on tasks of a different type (say, type B). Clark and Mayer (2011) provide a similar recommendation suggesting that practice (with feedback) should be designed to build job-relevant skills. A straightforward application of this theory, leads to the hypothesis that children in the Guided Discovery condition (GD) should perform better on the paper pre and post-tests than those in other conditions, because GD is better aligned with the goals of the tasks on the paper pre/post tests. The tasks on the paper test assess children's ability to predict, observe and explain (those are the "job-relevant" skills in Clark & Mayer's terms) and the GD condition involves practice on predicting, observing, and explaining. On the other hand, children in the Exploration (E) condition, by this theory, should perform better than child in other conditions on the Tower Building pre/post tests, since E provides the most practice building their own towers. The theory here is simple: if you practice tower building more, you learn tower building better. If you practice predicting and explanation more, you get better at predicting and explaining what will happen. This principle does not make strong predictions about differences in enjoyment in the different conditions.

A similar theory puts less focus on similarity in task types and more emphasis on similarity in underlying knowledge (facts, concepts, skills, principles) acquired in learning and transferred to assessments (cf., Koedinger, Perfetti, & Corbett, 2013). It also suggests that guidance is critical to effective acquisition. A consequential prediction is that guided discovery will produce better learning of the fundamental physics principle and those principles are needed (and will transfer to) the building task. Although unguided or minimally guided instructional approaches are very popular and intuitively appealing, there is evidence from empirical studies that indicate that minimally guided instruction is less effective and less efficient than instructional approaches that place a strong emphasis on guidance of the student learning process (Kirschner et al., 2010). Thus, by this theory, the Guided Discovery (GD) condition show better learning than Exploration (E) and Combined (GDE) conditions on both the paper pre/post tests and the transfer tower building activity. According to this view, principles of balance are better acquired in GD condition, and are needed for effective tower building. In line of this argument, there is an assumption that these acquired principles transfer easily to tower building. There is evidence that guided-discovery learning transfers to tower building from my previous experiments reported above (Yannier et al., 2015). However there is still room for improvement. Also, in this view, GD and GDE conditions may have greater enjoyment compared to the E condition, as in the Exploration condition children may have frustration at failure since they don't get any guidance.

A third line of theoretical argument is based on constructivist theory: that humans generate knowledge and meaning from an interaction between their experiences and their ideas while exploring freely without explicit instruction. This view is especially prevalent in informal learning settings: many museum exhibits are designed to support mere exploration with physical materials (Jeffery-Clay, 1998). Also, many researchers claim that tangible interfaces and mixed-reality environments are well suited for exploratory activity (Marshall, 2007). According to this view, the Exploration condition would perform better on both the paper pre/post tests, tower building activities and enjoyment measures (survey).

A final fourth line of theoretical argument recognizes that guided-discovery and exploration may have complementary benefits within a mixed-reality environment – it is essentially a combination of the first and second. This view is supported by the insights I have gained through the previous studies. We have seen that the guided-discovery method (with the predict observe explain structure supported by interactive feedback) has lead to high learning gains. We have also seen that children use mental visualizations when they observe physical phenomena, which leads to better learning. The addition of more exploratory activities where they get to build their own towers may enhance their mental visualizations and embodied cognition, thus leading to better learning (especially in transfer to tower building tasks). Also, we have seen that kids enjoy the open-ended building activities more. Thus, the combination of guided discovery and exploration may have complementary benefits leading to better learning and enjoyment. If this hypothesis is true, we would expect students in the Combined condition to perform best in tower building tasks, enjoyment measures (may be same as E) and paper pre/post tests (may be same as GD). Guided-discovery activities would prepare the student for the exploration activities, thus leading to more learning when combined. The Combined condition would also lead to more agency and self-efficacy since the guided discovery produces more effective performance on exploration, thus leading to more enjoyment. Frustration about GD tasks may block agency on the child's part leading to less enjoyment, where as the Exploratory building tasks may be inherently more enjoyable compared to GD.

	Paper Pre/post tests	Tower building	Enjoyment
You learn what you practice	GD>GDE>E	E>GDE>GD	? GD=GDE=E
Guidance is better for fundamental principle learning	GD>GDE>E	GD>GDE>E	GD=GDE>E
Constructivist/Exploration is better for learning	E>GDE>GD	E>GDE>GD	E>GDE>GD
Complementary benefits	GDE=>GD>E	GDE>GD>E	GDE>E>GD

Table 3. Hypotheses based on theories or learning principles.

8.4 Design of Mixed-Reality Educational System NoRILLA: Novel Research-based Intelligent Lifelong Learning Apparatus

I have done some interviews with teachers asking them what they think about our mixed-reality system. Most of the feedback was very positive: expressing that they think this system could help the teachers as well as their students learn science and inquiry together in a much more interactive and collaborative way. One of the teachers said:

"The more the video and screen generation comes through, the shorter their attention span is. I feel like I'm competing with the Xbox, the Wii. I have to be super engaging for them to pay attention to me. There is so much technology out there for kids, that's great but there is so few ways to get them on the same thing at the same time. I love that NoRILLA uses technology in such an engaging, communicative and non-isolating way. I'm not a scientist, I'm not a scientist by any stretch of imagination and I love science and I love to teach science, but I feel like I'm limited by own limitations in the science world. To have something like this that supports and backs up and lets the kids and myself all learn together is genius!"

A common comment that most teachers pointed out was that they really liked the way our system encourages productive dialogue by bridging physical and virtual worlds and they wished that this could be extended to many different content areas that they teach, creating a science curriculum around it.

Thus, I will work with some teachers to create new activities and science curriculum that NoRILLA can be extended to. Some possible content areas include Balancing and Weighing, Forces and Motion, Planetary Systems, Density, Weather, Human Body and Simple Machines. åSome possible future designs may include Balance Scale that is integrated with a game that teaches kids the principles of balance and moment of inertia, and Ramps and Cars that is integrated with a game that teaches about the principles of Forces and Motion.

9. Timeline

October 15-30	Thesis proposal presentation
November to December	Make modifications based on museum pilot, get the program running
January to February	Pilot testing of timing and logging
January to March	Build different conditions
April to May	Run experiment
May to July	Data Analyses
July to September	Write the Thesis
September/October	Defend

10. Conclusions and Contributions

With this work, I aim to make different contributions:

First, my work demonstrates that a mixed-reality environment bridging physical and virtual worlds can improve children's learning and enjoyment above and beyond equivalent tablet or computer versions. It also offers insight as to why experimenting in the real world with interactive feedback may improve children's learning, providing evidence about gestures that they use as signs of their mental visualizations.

Second, I aim to investigate how to create a mixed-reality environment that is optimized for learning and enjoyment. What are the features that lead to more learning in a mixed-reality game (guided-discovery versus exploration/problem solving versus a combination of both)? Which of these features make it more enjoyable? My contribution from a learning science perspective will be do discover which features make the mixed-reality environment more educational and enjoyable and make recommendations accordingly.

Lastly, my work aims to make a contribution from the design perspective by creating a new educational system that bridges physical and virtual environments to improve children's learning and enjoyment in a collaborative way, fostering productive dialogue and scientific curiosity to be used in museum, school and home settings. My work will also include some design features, recommendations and scenarios that can be extended to different content areas and inform early science learning for young children.

11. References

Abrahamson, D., Lee, R. G., Negrete, A. G., & Gutiérrez, J. F. (2014). Coordinating visualizations of polysemous action: values added for grounding proportion. *ZDM*, *46*(1), 79-93.

Aleven, V., Koedinger, K. R. The need for tutorial dialog to support self-explanation. *Building dialogue systems for tutorial applications, papers of the 2000 AAAI Fall Symposium*. 2000.

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning?. *Journal of Educational Psychology*, *103*(1), 1.

Antle, A. N. (2012). Research opportunities: Embodied child–computer interaction. *International Journal of Child-Computer Interaction*.

Antle, A. N., Droumeva, M., & Ha, D. (2009, June). Hands on what?: comparing children's mouse-based and tangible-based interaction. *Proceedings of the 8th International Conference on Interaction Design and Children* (pp. 80-88)

Azmitia, M., & Crowley, K. (2001). The rhythms of scientific thinking: A study of collaboration in an earthquake microworld. Designing for science: Implications from everyday, classroom, and professional settings.

Back, M, Cohen, J, Gold, R, Harrison, S and Minneman,S (2001).Listen Reader: an electronically augmented paper-based book. Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'01), ACM Press, pp23-29

Bakker, S., van den Hoven, E., & Antle, A. N. (2011, January). MoSo tangibles: evaluating embodied learning. In *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*(pp. 85-92).

Benko, H., Jota, R., and Wilson, A. MirageTable: Freehand Interaction on a Projected Augmented Reality Tabletop. *CHI 2012.*

Carini, R. M., George D. K., and Stephen P. Klein. Student engagement and student learning: Testing the linkages*. *Research in Higher Education* 47.1 (2006): 1-32.

Center for the Digital Future. The 2009 digital future report: Surveying the digital future-year eight. Los Angeles: USC Annenberg School Center for the Digital Future (2009)

Chase, C.C., Shemwell, J.T., & Schwartz, D.L. (2010). Explaining across contrasting cases for deep understanding in science: An example using interactive simulations. Proceedings of the 2010 International Conference of the Learning Sciences.

Christel, M. G., Stevens, S. M., Maher, B. S., Brice, S., Champer, M., Jayapalan, L., ... & Lomas, D. (2012, July). RumbleBlocks: Teaching science concepts to young children through a Unity game. In *Computer Games (CGAMES), 2012 17th International Conference on* (pp. 162-166). IEEE.

Clark, Ruth C., and Richard E. Mayer. *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning*. John Wiley & Sons, 2011.

Cook, S., Wagner, Z. M., and Goldin-Meadow, S. Gesturing makes learning last. *Cognition* 106.2 (2008): 1047-1058.

Corbett, A. T., and Anderson, J. R. Locus of feedback control in computer-based tutoring: Impact on learning rate, achievement and attitudes. *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2001.

Glenberg, Arthur M. "What memory is for: Creating meaning in the service of action." Behavioral and brain sciences 20.01 (1997): 41-50.

Gokhale, Anuradha A. "Collaborative learning enhances critical thinking." (1995).

Gupta, A., et al. DuploTrack: a real-time system for authoring and guiding duplo block assembly. *Proceedings of the 25th annual ACM symposium on User interface software and technology*. ACM, 2012.

Harvel, C. (2010). Guided discovery learning. *Faithbased education that constructs: a creative dialogue between constructivism and faith-based education. Wipf and Stock Publishers, Eugene*, 169-172.

Hayne, H., Herbert, J. & Simcock, G. (2003). Imitation from television by 24- and 30-month-olds. Developmental Science, 6(3), 254-261.

Henning, P. (1998). Everyday Cognition and Situated Learning. In Jonassen, D. (Ed.), Handbook of Research on Educational Communications and Technology. (2nd. Ed.). New York: Simon & Schuster

Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. Psychonomic bulletin & review, 15(3), 495-514.

Jeffery-Clay, K. R. (1998). Constructivism in museums: How museums create meaningful learning environments. *The Journal of Museum Education*, 3-7.

Kirschner, Paul A., John Sweller, and Richard E. Clark. "Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching."*Educational psychologist* 41.2 (2006): 75-86.

Klahr, D., Triona, L. M., and Williams, C. Hands on what? The relative effectiveness of physical versus virtual materials in an engineering design project by middle school children. *Journal of Research in Science Teaching* 44.1 (2007): 183-203.

Koedinger, K.R., McLaughlin, E. A., & Heffernan, N. T. (2010). A quasi-experimental evaluation of an on-line formative assessment and tutoring system. /Journal of Educational Computing Research, 43 /(4), 489-510.

Kozulin, A., Gindis, B., Ageyev, V. S., & Miller, S. M. (Eds.). (2003). *Vygotsky's Educational Theory in Cultural Context. Learning in Doing: Social, Cognitive, and Computational Perspectives*. Cambridge University Press.

Lipsey, M.W., Puzio, K., Yun, C., Hebert, M.A., Steinka-Fry, K., Cole, M.W., Roberts, M., Anthony, K.S., Busick, M.D. (2012). Translating the Statistical Representation of the Effects of Education Interventions into More Readily Interpretable Forms. (NCSER 2013-3000). Washington, DC: National Center for

Special Education Research, Institute of Education Sciences, U.S. Department of Education. This report is available on the IES website at http://ies.ed.gov/ncser/.

Marshall, P. (2007, February). Do tangible interfaces enhance learning?. In*Proceedings of the 1st international conference on Tangible and embedded interaction* (pp. 163-170). ACM.

Martin, T., & Schwartz, D. L. (2005). Physically distributed learning: Adapting and reinterpreting physical environments in the development of the fraction concept. Cognitive Science, 29, 587–625

Montessori, M. *Montessori method*. Random House Digital, Inc., 1964.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, *2*(2), 175.

O'Malley, C., Stanton-Fraser, D.: Literature review in learning with tangible technologies, Nesta FutureLab Series, report 12 (2004).

Olympiou, G., and Zacharias C. Z. Blending physical and virtual manipulatives: An effort to improve students' conceptual understanding through science laboratory experimentation. *Science Education* 96.1 (2012): 21-47.

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books, Inc..

Puchner, L., Rapoport, R., & Gaskins, S. (2001). Learning in children's museums: is it really happening?. *Curator: The Museum Journal*, *44*(3), 237-259.

Quinn, H. et al., A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas, The National Academies Press (2012)

Resnick, M, Maryin, F, Berg, R, Boovoy, R, Colella, V, Kramer, K et al (1998). Digital manipulatives: new toys to think with. Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, 281-287

Rieser, J.J., Garing, A.E. and Young M.F.(1994). Imagery, action and young children's spatial orientation- it's not being there that counts, it's what one has in mind. Child Development, 65(5), 1262- 1278

Roe, K., Mujis, D., Children and computer games – a profile of the heavy user. European Journal of Communication, 13(2), pp. 181-200 (1998)

Ryokai, R. and Cassell, J. (1999). StoryMat: a play space for collaborative storytelling. Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, Pittsburgh, PA, ACM Press, pp272-273

Schneider, Bertrand, Engin Bumbacher, and Paulo Blikstein. "Discovery versus direct instruction: Learning outcomes of two pedagogical models using tangible interfaces." *Exploring the material conditions of learning: opportunities and challenges for CSCL, B the proceedings of the Computer Supported Collaborative Learning (CSCL) conference. Gothenburg: ISLS*. 2015.

Schneider, B., Jermann, P., Zufferey, G., & Dillenbourg, P. (2011). Benefits of a Tangible Interface for Collaborative Learning and Interaction. *IEEE Transactions. on Learning Technologies*, 4, 222–232.

Shelley, T., Lyons, L., Zellner, M., & Minor, E. (2011, May). Evaluating the embodiment benefits of a paper-based tui for educational simulations. In*CHI'11 Extended Abstracts on Human Factors in Computing Systems* (pp. 1375-1380).

Stanton, D., O'Malley, C., Ng, K. H., Fraser, M. and Benford, S. (2003 (June)). Situating historical events through mixed reality: adult child interactions in the storytent, in: SLB Wasson, U Hoppe (Ed), Designing for Change in Networked Learning Environments (Vol 2, pp293-302)

Tseng, T., Bryant, C., & Blikstein, P. (2011, June). Collaboration through documentation: automated capturing of tangible constructions to support engineering design. In Proceedings of the 10th International Conference on Interaction Design and Children (pp. 118-126). ACM.

Tuovinen, J. E., & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of educational psychology*, *91*(2), 334.

Turkle, S. (2015). Reclaiming Conversation: The Power of Talk in a Digital Age.

Ullmer, B. and Ishii, H. Emerging frameworks for tangible user interfaces. IBM Syst. J. 39, 3-4 (2000), 915–931.

Underkoffler, J. and Ishii, H., Illuminating light: an optical design tool with a luminous-tangible interface. In Proc. of CHI '98, ACM Press, 542-549.

Vygotsky, Lev. "Interaction between learning and development." Readings on the development of children 23.3 (1978): 34-41.

Walker, E., & Burleson, W. (2012, May). Using need validation to design an intelligent tangible learning environment. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems* (pp. 2123-2128). ACM.

Yannier, N., Basdogan, C., Tasiran, S., Sen, O. L., 2009, Using Haptics to Convey Cause and Effect Relations in Climate Visualization, IEEE Transactions on Haptics, Vol. 1, No. 2, pp.130-141

Yannier, N., Koedinger, K. R. and Hudson, S. E. Tangible Collaborative Learning with a Mixed-Reality Game: EarthShake. *Artificial Intelligence in Education*. Springer Berlin Heidelberg, 2013.

Yannier, N., Koedinger, K. R., & Hudson, S. E. (2015, April). Learning from Mixed-Reality Games: Is Shaking a Tablet as Effective as Physical Observation?. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 1045-1054). ACM.