Adding Physical Objects to an Interactive Game Improves Learning and Enjoyment: Evidence from EarthShake

NESRA YANNIER and SCOTT E. HUDSON, Carnegie Mellon University
ELIANE STAMPFER WIESE, University of California, Berkeley
KENNETH R. KOEDINGER, Carnegie Mellon University

Can experimenting with three-dimensional (3D) physical objects in mixed-reality environments produce better learning and enjoyment than flat-screen two-dimensional (2D) interaction? We explored this question with EarthShake: a mixed-reality game bridging physical and virtual worlds via depth-camera sensing, designed to help children learn basic physics principles. In this paper, we report on a controlled experiment with 67 children, 4–8 years old, that examines the effect of observing physical phenomena and collaboration (pairs vs. solo). A follow-up experiment with 92 children tests whether adding simple physical control, such as shaking a tablet, improves learning and enjoyment. Our results indicate that observing physical phenomena in the context of a mixed-reality game leads to significantly more learning and enjoyment compared to screen-only versions. However, there were no significant effects of adding simple physical control or having students play in pairs vs. alone. These results and our gesture analysis provide evidence that children's science learning can be enhanced through experiencing physical phenomena in a mixed-reality environment.

Categories and Subject Descriptors: H.5.m. [**Information Interfaces and Presentation (e.g., HCI)**]: Miscellaneous

General Terms: Human Factors

Additional Key Words and Phrases: Tangible interfaces, learning technologies, mixed-reality learning, educational games

ACM Reference Format:
Nesra Yannier, Scott E. Hudson, Eliane Stampfer Wiese, and Kenneth R. Koedinger. 2016. Adding physical objects to an interactive game improves learning and enjoyment: Evidence from earthshake. ACM Trans. Comput.-Hum. Interact. 23, 4, Article 26 (September 2016), 31 pages.
DOI: http://dx.doi.org/10.1145/2934668

1. INTRODUCTION

Today's children are drawn into the compelling world of two-dimensional (2D) flat-screen technologies, such as tablets or computer games, starting from early childhood.

The research is partly supported by contract ONR N00014-12-C-0284 (awarded by DARPA), "Learning to Solve Problems, Solving Problems to Learn", partly by LearnLab (NSF grant SBE-0836012) and partly by Carnegie Mellon University's Program in Interdisciplinary Education Research (PIER) funded by Grant R305B090023 from the U.S. Department of Education, and by the Institute of Education Sciences, U.S. Department of Education, through Grant R305C100024 to WestEd. Any opinions expressed in this paper represent those of the authors, not ONR, DARPA, NSF, the Institute or the U.S. Department of Education. We would also like to thank the participating schools, teachers and students.
Authors' addresses: N. Yannier and S. E. Hudson, Human Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, PA 15213; emails: nyannier@cs.cmu.edu, scott.hudson@cs.cmu.edu; E. S. Wiese, University of California, Berkeley, Berkeley, CA; email: eliane.stampfer.wiese@gmail.com; K. R. Koedinger, Human Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, PA 15213; email: koedinger@cmu.edu.

DOI: http://dx.doi.org/10.1145/2934668

Screen-based technologies can help children learn by providing immediate targeted feedback

[Corbett and Anderson 2001]. However, as screen-based technologies are becoming more appealing for children, it is worth asking whether real world interaction is still needed to enhance learning and enjoyment. Are today's children, immersed in 2D flat-screen technologies, missing out on opportunities in their three-dimensional (3D) physical environment, where they may develop understanding more readily? Some have argued that these flat-screen technologies can have negative effects on children [Roe and Mujis 1998; Lebo 2007]. Screen-based technologies also have a tendency to pull people away from their physical environment and make them physically and socially isolated. Roe and Mujis have found some justification to associate frequent gamers with social isolation and less positive behavior towards society [Roe and Mujis 1998]. Researchers at University of Southern California have shown that family time has decreased by more than 30% due to computer usage at home [Lebo 2007].

How can games for young children combine the distinct advantages of screen-based technologies and the 3D physical world? Screen-based educational games can provide students with instructional support, such as immediate correctness feedback and instructional prompts [Corbett and Anderson 2001]. Screen-based games also have motivational benefits, such as compelling scenarios and engaging characters. Screen-based interactive visualizations often aid learning, comprehension, and thinking by making abstract concepts more visible, especially in science [Uttal and Doherty 2008]. On the other hand, most intuitive learning occurs in our physical 3D world, arguably where learning is at its best [Henning 2004]. The physical world is a child's everyday environment, where she plays, discovers, experiments, and learns – often collaboratively with others. 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 and 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 [Ullmer and Ishii 2000] and mixed-reality environments [Rogers et al. 2002] have the potential to help students learn in more engaging and powerful ways than either modality alone.

Although there are many compelling tangible interfaces, there are too few experimental tests of the hypothesis that including real, 3D objects may improve student learning [Rieser et al. 1994]. Furthermore, we do not have a sufficient empirical basis for evaluating alternative explanations for why and how 3D physical objects may enhance learning compared to 2D flat-screen representations. Most previous studies do not identify what it is that provides benefits for learning in these mixed-reality environments: do students benefit from physical triggers, or is it the observations of the resulting physical phenomena that primarily drive learning? For example, in Listen-Reader [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. 1998a], designed to help students learn principles of acceleration, are learning benefits derived from the action of throwing the ball, or rather from observing a physical 3D ball rather than a virtual 2D one? Across tangible interfaces, some forms of hands-on manipulation are more related to the learning context (i.e., the experience of throwing BitBall forcefully or softly may help students make sense of the resulting acceleration measurements), while others are less so (i.e., the action of turning a page may not directly support reading comprehension). To what extent do students benefit from observing real, physical objects, beyond the benefits of seeing those objects portrayed on a screen? To what extent do students benefit from physical controls that are not directly related to the learning goals but may increase enjoyment? Through rigorous controlled experimentation, we examine if observing physical phenomena in a mixed-reality environment can improve learning and enjoyment beyond that of a screen-only control. We further examine if a simple physical trigger can provide benefits beyond those of a mouse-only control.

The following sections describe prior work and provide theoretical background. First, we 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 learning with 3D physical objects useful, and under what conditions. Next we 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 that 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 and Burleson 2012]. This paper is an attempt to begin to answer the questions left by both literatures.

1.1. Background

1.1.1. Learning with Physical Objects and 3D Representations. Experiments on the role of physical objects in learning have produced mixed results. We first present research that found benefits for 3D physical objects/representations over 2D representations of the same concepts. Children learn less from 2D representations on television than they do from 3D objects in live demonstrations, termed the video deficit effect [Barr 2010]. Further, 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]. Spatial reasoning, which is particularly important for STEM learning, may be enhanced with 3D physical objects [Davis 2015]. Martin and Schwartz showed that manipulating physical chips facilitated children's interpretation of fractions better than seeing an image of the grouped pieces on paper, though they only compared performance with these scaffolds, not learning after the scaffolds were removed [Martin and Schwartz 2005]. Gilbert pointed out the importance of types of representation and the modes in which they are expressed (e.g., concrete, material mode of representation that retains the three dimensions of that which is being represented) [Fitzmaurice et al. 1995]. Additionally, Manches created a framework about the benefits of physical manipulatives, stressing the advantages of physical action for conveying information, activating real-world knowledge, and improving memory [Manches 2011]. Physical manipulatives may provide benefits by being an additional channel to convey information, activating real-world knowledge, and improving memory through physical action [Manches and Price 2011].

Other research has demonstrated no added learning benefit of physical materials over virtual analogs. Klahr et al. found no differences for middle school students' learning of experimental design principles when setting up experiments with physical vs. virtual springs [Klahr et al. 2007]. Marshall et al. compared learning from a balance task with physical vs. graphical materials and found no difference for adults (ages 18–46) [Marshall et al. 2010]. These experiments did not include any interactive feedback and did not examine mixed-reality conditions; students interacted with either physical or virtual materials on their own. In another experiment in the context of light and color, Olympiou and Zacharias also found no difference in learning from only physical vs. only virtual materials for university students [Olympiou and Zacharia 2012]. 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 and Zacharia 2012].

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

1.1.2. 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 [Rieser et al. 1994]. In tangible interfaces, a person interacts with digital information through the physical environment. Similarly, mixed-reality environments combine physical and virtual worlds in an interactive way. These environments hold promise for education in that they can provide the benefits of physical objects while giving feedback and other instructional support to students. Some researchers have investigated the potential of tangible environments for supporting collaborative learning [Falcão and Price 2009]. Others have analyzed the importance of embodied interaction and discourses of scientific investigation using an interactive tangible tabletop [Valdes et al. 2012]. Hornecker and Buur have created a framework on physical space and social interaction for tangible interfaces, stressing that our understanding of human interaction with hybrid or augmented environments is very limited [Hornecker and Buur 2006]. Price et al. have also investigated the role of embodied cognition with tangibles for learning [Price et al. 2009].

Many researchers have instrumented objects for learning to make them interactive (for example, 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 and Cassell 1999]; a ball that measures and shows its acceleration [Resnick et al. 1998b]; a mixed-reality experience that helps children discover and reflect on historical places and events [Stanton et al. 2003]; an interactive display for children to create, record, view, and test systems of tangible simple machine components [Tseng et al. 2011]; and Fabulous Beasts: a game of stacking smart objects for two players (http://playfabulousbeasts.com/). However, many tangible interfaces were studied as prompts for student investigation and exploration – that is, the researchers wanted to see how students would use these objects without instructions. Therefore, this body of work does not address the role of physical objects in learning: the objects were not compared to a control and the experiments did not have post-test assessments of learning. In contrast, instead of designing an interaction for pure exploration, our goal is to create a mixed-reality game with guided feedback and self-explanation, as these two pedagogical supports have been shown to enhance learning [Corbett and Anderson 2001; Aleven and Koedinger 2000]. We aim to augment the physical environment with synchronized, interactive feedback, and inquiry-based activities to produce a pedagogically strong and engaging learning experience. Additionally, we aim to determine the effects of observing physical objects 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 for manipulating the pitch, volume, and tempo of ongoing tones [Back et al. 2001]. Their qualitative interviews and video analysis indicate that MoSo provided children with a physical handle to reason about the targeted abstract sound concepts [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 [Ryokai and Cassell 1999]. 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 seeing a virtual environment [Yannier et al. 2008].

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: through observing phenomena in the physical environment, through physical triggers that make them more enjoyable, or through manipulating physical objects. To untangle the effects of each, we need randomized controlled experiments that isolate these variables.

1.2. 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) *mental visualizations*: experiencing physical phenomena in the real, 3D world facilitates mental visualizations and cues analogs to reason with; (2) *enjoyment*: physical experiences are inherently more enjoyable; and (3) *collaboration:* the physical environment provides more opportunities for collaboration, which enhances learning. We discuss each in turn.

First, experiencing a physical phenomenon with 3D representations of physical objects in the real world may help people perceive and mentally visualize the target objects [Antle 2013; Gokhale 1995; Engelkamp and Zimmer 1989], 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 processed more deeply, allowing for recognition of 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, which suggests that when children (and adults) learn or reason with abstract concepts, they utilize mental simulations based on concrete motor–perceptual experiences [Antle 2013]. Alibali et al. have theorized that perceptual and motor simulations underlie embodied language and mental imagery, and are often revealed by spontaneous gestures that accompany speech [Hostetter and 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 re-run, re-activating the same neural patterns [Montessori 1964]. 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. 2014]. 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,

Fig. 1. Students interacting with EarthShake. Left: an example of the mixed-reality and pair condition of the experiment. Right: an example of the virtual and solo condition.

observing real-world phenomena in a mixed-reality environment may trigger mental simulations and affordances for action, facilitating later 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 may 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 physical environments [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 physical objects may improve learning.

Thus, adding physical objects to an interactive game might improve learning for children. To test this hypothesis, we designed two carefully controlled experiments comparing learning outcomes within a simple interactive game with guided feedback. In the first experiment, we 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 we compared students playing in pairs to students playing solo. In the second experiment, we again compared the mixed-reality versions of EarthShake with equivalent screen-based versions. However this time we also added a potentially engaging simple physical control (such as shaking the tablet to create the earthquake on the screen). The second experiment investigates if adding an inherently more enjoyable physical/hands-on control can increase learning by increasing enjoyment or if physical observation and experimentation are more critical. Below we review EarthShake and our experiments in more detail.

2. EARTHSHAKE

EarthShake (see Figure 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 (e.g., center of mass, wide base, height, and symmetry) [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 (experimenting, discovering, and learning with physical objects, and facilitated face-to-face social interaction and collaboration).

As shown in Figure 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. After a simulated earthquake shakes the table, the system detects which of the towers in the physical setup fell first 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 Vygotsky's theories of learning by doing and minimal assistance [Vygotsky 1978]. According to these theories, instruction should target a student's Zone of Proximal Development, where students cannot achieve the given task independently, but can do so with guidance [Vygotsky 1978]. In EarthShake, children are not directly told about the physics principles (symmetry, center of mass, wide base, height, and so on) 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 and pedagogical prompts on their predictions. To further facilitate mental construction of these key ideas, EarthShake uses prompted self-explanation [Aleven and Koedinger 2000].

EarthShake is designed for children age 4–8 (K-3 grade) to engage them in STEM early on. Early engagement is important: the National Center for STEM Elementary education found that a third of US elementary students have lost interest in science by fourth grade.[1] EarthShake aims to teach students principles of stability and balance, which are listed in the United States' National Research Council (NRC) Framework & Asset Science Curriculum for this age group [Quinn et al. 2012]. Teacher interviews and a cognitive task analysis identified the scientific thinking that is needed to be successful with stability and balance [Christel et al. 2012]. A key result was the explication of four principles of physics: 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). These principles were used to develop contrasting cases, an instructional approach that presents two examples that only differ on one important feature. Contrasting cases help novices focus on those important domain features and have been shown to be beneficial for deep understanding in science [Chase et al. 2010]. EarthShake also builds on Azmitia and Crowley's research, which stresses the importance of scientific thinking and collaboration in an earthquake micro-world [Azmitia and Crowley 2001].

2.1. Scenario

Here, we describe the mixed-reality version of EarthShake, as played in pairs. EarthShake is structured around a predict/observe/explain cycle. The game starts with a gorilla character asking the pair of students which of the two towers will fall first when he shakes the table [Yannier et al. 2013]. The students 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, the students

[1] http://www.usnews.com/news/articles/2011/08/29/stem-education–its-elementary.

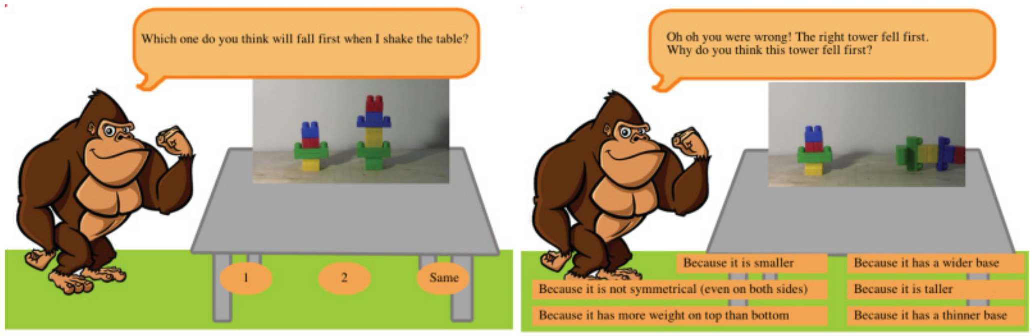

Fig. 2. Virtual-only version of Earthshake, showing the predict/observe/explain cycle. The video of the physical towers shaking on the earthquake table is integrated into the game interface.

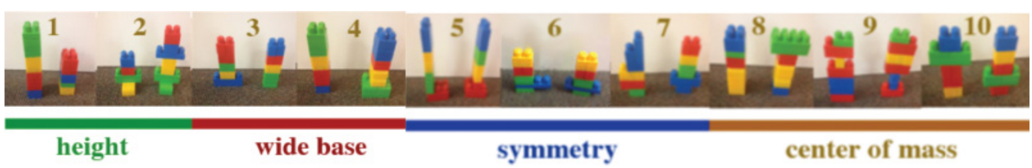

Fig. 3. Contrasting cases used during the game.

use a mouse to click on the virtual representation of the tower that they predict will fall first. The gorilla then tells the pair to discuss why that 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 falls. 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. 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,” and “Because it has a thinner base” (Figure 2). This scenario is repeated for different contrasting cases targeting the height, wide base, symmetry, and center of mass principles (Figure 3). Note that while students observe the physical towers, they do not touch them.

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 4). The Kinect camera detects when a tower falls, ensuring that EarthShake is in sync with the physical 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

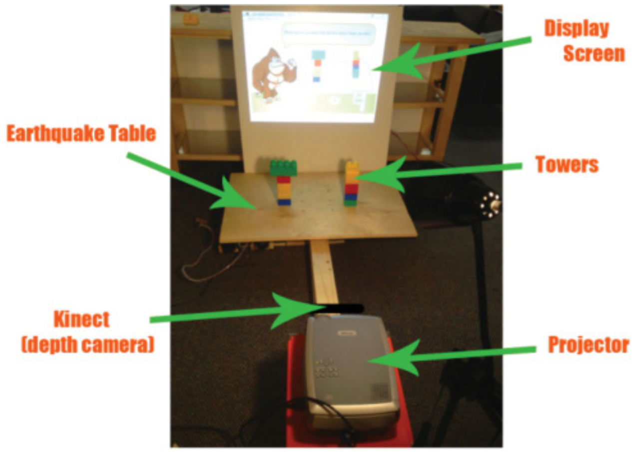

Fig. 4. Physical setup of EarthShake.

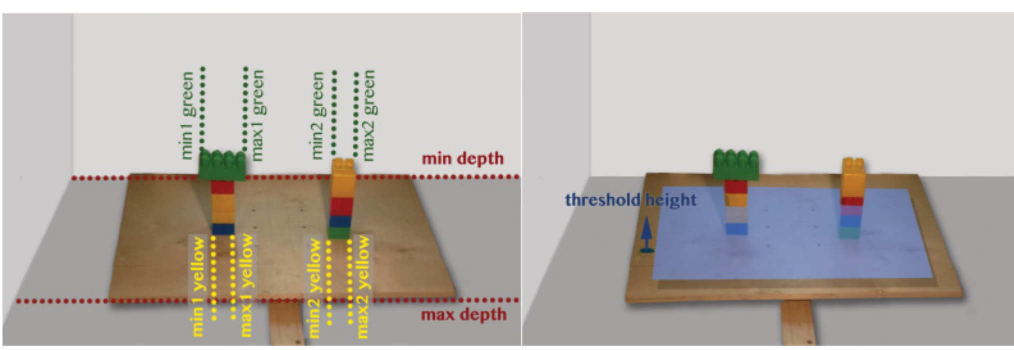

Fig. 5. Our vision algorithm first calculates the minimum and maximum values for each color blob in each tower to determine where each tower stands (top). Then we use a threshold height to detect a fall, if all the color blobs in the tower are below this height (bottom).

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 (Figure 5).

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 and 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. We use the Kinect camera and our 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.

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 females; grades K-3). 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 pre-test, and 78% answered correctly at post-test ($t(11) = 4.2$, $p < 0.002$, $d = 0.78$). On items asking students to explain why a tower fell first, 17% answered correctly at pre-test, whereas 71% answered correctly at post-test ($t(11) = 9$, $p < 0.001$, $d = 2.98$). 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]. Note that while students built towers before and after playing Earthshake, they did not build or even touch the towers in the game.

Qualitative video analysis indicated high levels of enjoyment and excitement when the table shook and made the towers fall. The children also had 'a-ha' moments after making incorrect 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 explained their reasoning to each other, with statements such as "Look! That one will fall first because it has a bigger top." These explanations continued after the game. In one particular instance of collaboration and joint explanation development, one child guided another in the tower-building task. When the first child started to put more blocks on one side of his tower than the other, 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].

We designed the next 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 the observation of physical objects coupled with interactive feedback might play an important role, as it seemed to increase enjoyment and embodied cognition – for example some children mentioned that they believed seeing what happens in real life more rather than having the computer tell them what would happen. Additionally, students' collaborations and discussions might have enhanced their learning [Yannier et al. 2013]. To separate the factors of media type and collaboration, we designed a 2×2 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 enjoyment and/or 3D mental visualization, which then yields greater learning, we would expect better learning from EarthShake for both solo and pair groups. We include measures of enjoyment and mental visualizations to evaluate their potential roles in mediating learning.

4. EXPERIMENT 1: MIXED-REALITY AND COLLABORATION

Our first experiment examines if observing physical objects leads to learning benefits beyond those provided by a matched video control. To do so, we compare 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 observing physical objects and directly manipulating them through touch, that is, sensory—motor interaction. 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-only or mixed-reality (physical with interactive feedback).

4.1. Experimental Method

4.1.1. Experimental Conditions. As illustrated in Figure 1, this 2×2 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, and 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. Students did not receive feedback on whether or not the selected explanations were correct. In the virtual condition, instead of watching physical towers fall, students observed pre-recorded videos. To make the conditions as equivalent as possible, we 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 in the virtual condition received the same feedback and explanation prompts as in the mixed-reality condition. Since the aim of the experiment was to isolate the learning benefits that could be attributed to observation alone, the method of interaction was held constant. 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 in both conditions were read aloud with voice over by the gorilla. 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.

4.1.2. Participants. 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 third grade (4–8 years old), 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. These participants came in to a lab for the study. The rest of the participants took part in the study at their school (two different local elementary schools with diverse student populations).

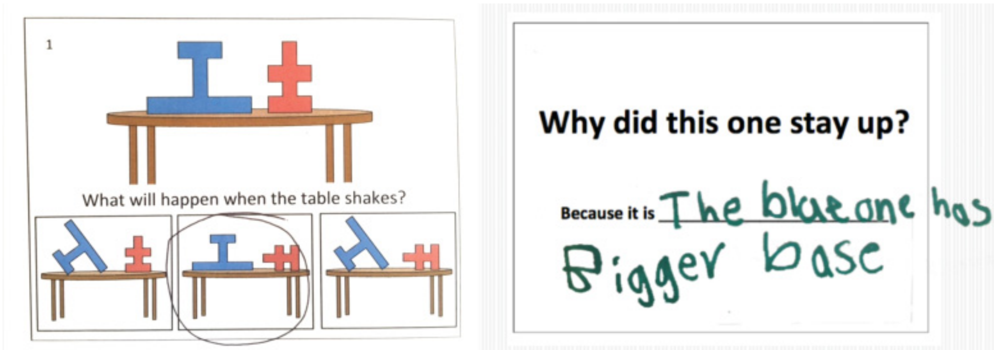

Fig. 6. Prediction (left) and explanation (right) items used in the paper pre-/post-tests.

Students assigned to the pair conditions were matched with siblings (in the lab setting) or classmates selected by their teachers (in the school setting). Students participating in the lab vs. at school were evenly distributed among the different conditions, as were pairs made up of siblings and classmates.

4.1.3. Procedure. Before playing, students independently completed a paper pre-test to measure what they already knew about the stability and balance principles in the game. The experimenter helped with reading the questions and writing down answers for the students who had difficulty reading or writing. Next, students did a tower-building task, which took approximately 3–5min. They were asked to use a given set of blocks to build a tower that would stay up when the earthquake table shook. 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, whereas 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 any changes in their towers after interacting with the game. After building the tower, they were given a matched paper post-test. Finally, the students took a one-question survey asking, "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," and "I liked it very much." The first author also briefly interviewed the participants to see what they liked and disliked about the activity and if they had any suggestions to improve it. The same procedure was used for both the virtual and mixed-reality conditions.

4.1.4. Measures. The paper pre- and post-tests were prepared based on the NRC Framework & Asset Science Curriculum [Quinn et al. 2012], and targeted the four principles of balance: symmetry, wide base, height, and center of mass. Questions presented a picture of two towers on a table, and asked students what would happen if the table shook. Prediction items asked students to select which tower would fall first, and explanation items asked for their reasoning (Figure 6). These questions were phrased as "What will happen when the table shakes?" and "Why did this one stay up?" Counting prediction and explanation items as individual questions, the tests had 21 questions in total.

4.2. Results of Experiment 1

This experimented investigated the effect of media type (mixed-reality vs. virtual) and collaboration (playing in pairs vs. solo) on students' learning and enjoyment. Learning was measured by the pre- and post-tests (both the paper assessments and tower-building tasks), whereas enjoyment was measured by the survey and qualitative observations. For our analyses, data from each child were treated as an independent observation, even for the children who played in pairs.

4.2.1. Paper-Test Results. We checked for differences at pre-test with a 2-way ANOVA on overall pre-test score and found no significant differences between the conditions (all F's < 0.79 and p's > 0.37). To test for learning, we ran a 2-way ANCOVA with post-test score as the outcome variable, pre-test as the covariate, and media type (mixed-reality vs. virtual) and collaboration (pair vs. solo) as fixed factors. We found a significant main effect of media type, with benefits for the mixed-reality condition. Collapsing across the collaboration conditions, average post-test scores were 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 main effect of collaboration and no interaction effect of media type and collaboration. These results show that the mixed-reality condition learned more than the virtual condition, both for students playing in pairs and playing solo (Figure 7(a)). There was no significant difference in time on task among the four conditions. Error bars in all the graphs represent standard error.

To investigate if the overall learning benefits for the mixed-reality conditions hold for both prediction items and explanation items, we analyzed each question type separately. To test for learning on the prediction items, we ran a 2-way ANCOVA with post-test prediction score as the outcome variable, pre-test prediction score as the covariate, and media type (mixed-reality vs. virtual) and collaboration (pair vs. solo) as fixed factors. There was a significant main effect of media type, with no significant effect for collaboration, and no significant interaction of media type and collaboration. The overall benefits of the mixed-reality condition held for the prediction items: the average post-test score for the mixed-reality condition was 76% while that of the virtual condition was 70% ($F(1,66) = 3.1$, $p < 0.0035$, $d = 0.39$). Note that chance is about 33% for these items (as most questions had three choices). These results show that the mixed-reality game improved prediction skills more than the virtual-only control, across both solo and pair conditions (Figure 7(b)). To test for learning on the explanation items, we ran a 2-way ANCOVA with post-test explanation score as the outcome variable, pre-test explanation score as the covariate, and media type (mixed-reality vs. virtual) and collaboration (pair vs. solo) as fixed factors. As with the prediction items, there was a significant main effect of media type, with no significant effect for collaboration, and no significant interaction of media type and collaboration (Figure 7(c)). The overall benefits of the mixed-reality condition held for the explanation items: the average post-test score for the mixed-reality condition was 52%, whereas that of the virtual-only condition was 26% ($F(1,66) = 18.6$, $p < 0.0001$, $d = 0.87$). Note the large effect size (0.87) of this difference in learning to provide a scientific explanation of stability.

We 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 (Figure 7(d) and (e). This finding that performance raises with grade level is evidence for the validity of our 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 percentage points per year whereas the effect of the mixed-reality condition over the virtual is 9.4 percentage

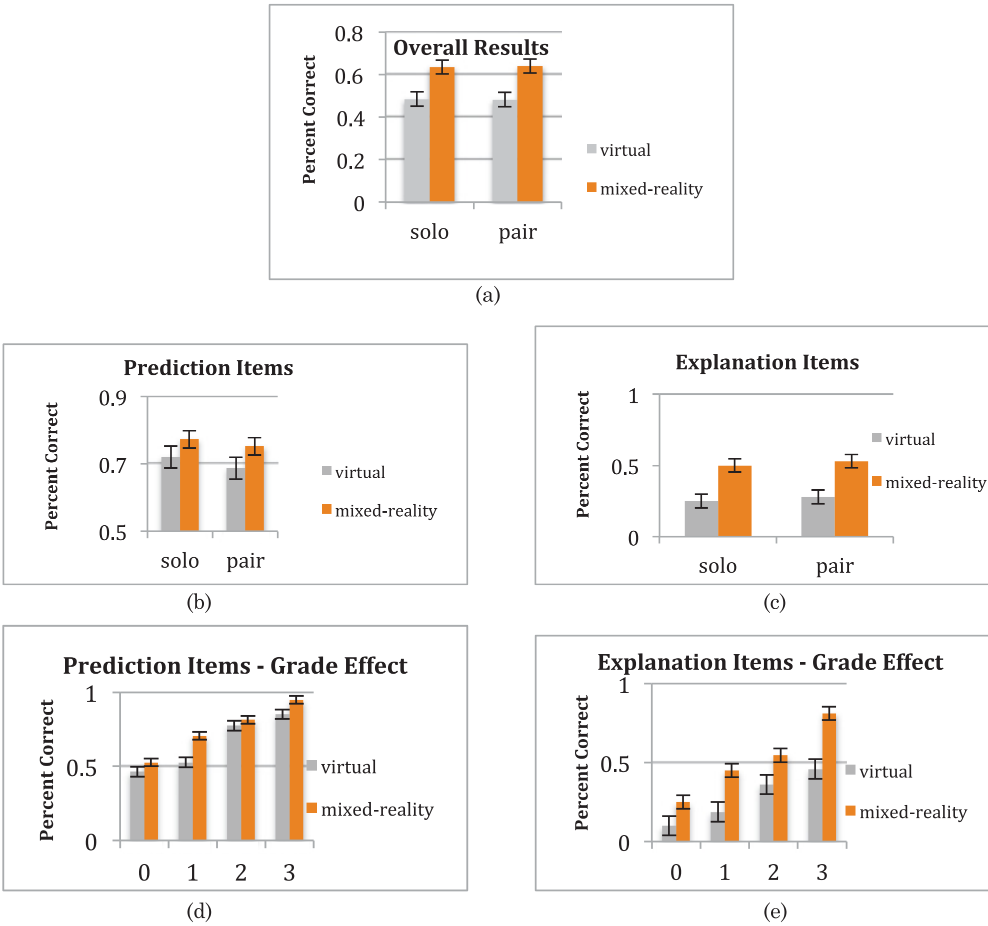

Fig. 7. Chance is 0.33 for these multiple-choice items. (a) Overall post-test scores. (b) Post-test scores for prediction items. (c) Post-test scores for explanation items. (d) Prediction scores by grade (K-3).

points.[2] Thus, this short treatment raises children's performance by about as much (9.4) as a year of normal schooling and maturation (9.5). This approach of using whole year increases as another way to gage the size of a treatment has been increasingly used [Koedinger et al. 2010] and recommended [Lipsey et al. 2012].

4.2.2. Tower-Test Results. To measure pre- to post-test changes on the tower-building task, we scored each student's towers according to three principles: height, symmetry, and center of mass (we 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

[2]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 second grader scores about 9.5 points higher than a first grader.

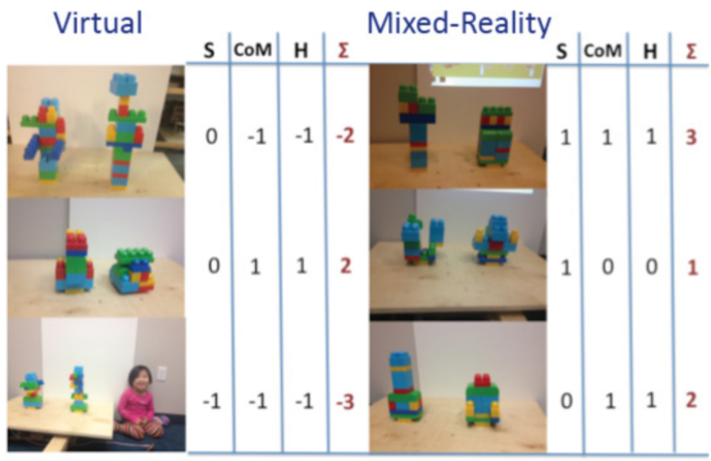

Fig. 8. Coding scheme for tower pre-/post-tests change.

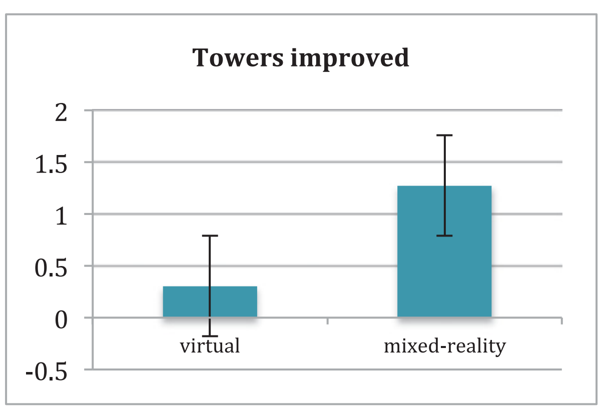

Fig. 9. Scores on the tower-building task (out of three). Positive scores indicate pre-to-post improvement in stability; a score of 0 indicates no change from pre to post.

those principles. Adding the scores for each principle yielded the student's total score (Figure 8).

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 9).

All three measures (the prediction items, the explanation items, and the towers) showed a significant positive effect of the mixed-reality conditions. What might explain this benefit? This paper explores three likely mechanisms suggested by prior work: collaboration, enjoyment, and mental visualizations [Carini et al. 2006; Antle 2013]. Comparisons of the solo and pair conditions did not suggest any effect of collaboration. Our quantitative and qualitative analyses, described below, provide evidence for mental visualizations but not enjoyment.

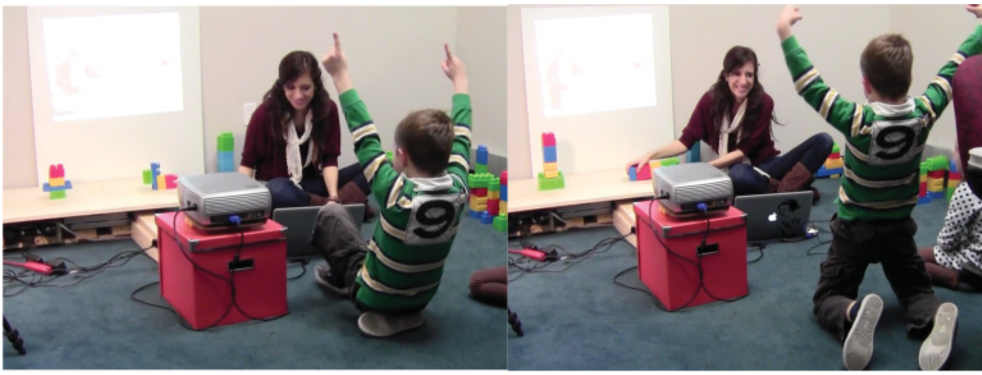

Fig. 10. A child exhibits enjoyment and excitement after a correct prediction.

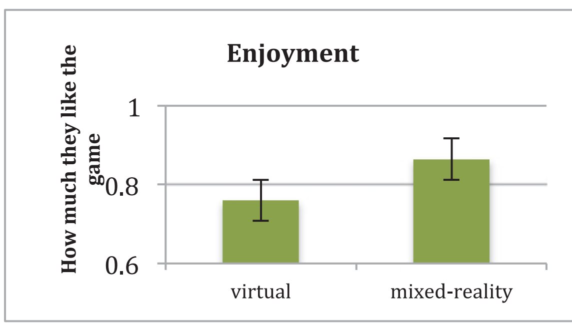

Fig. 11. Results of the survey given to measure how much children enjoyed the game.

4.3. 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 10). We did not see this level of enjoyment (e.g., jumping) in the virtual condition.

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 five-point Liker scale ("I didn't like it at all," "I didn't like it," "It was OK," "I liked it," and "I liked it very much"). Instead of matching a numeric score to each option, the scale used smiley faces to symbolize each emotion, so the children would better understand the choices in the scale. 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 (Figure 11 shows the mean ratings for likability, with the 1–5 scale converted to a proportion between 0 and 1). Although it seems reasonable that increased enjoyment may improve learning, we did not find evidence supporting the idea that increased enjoyment produces increased learning. There was not a strong correlation between enjoyment results and pre–post gains ($r = 0.21$). When we limit the learning analysis only to students who gave the highest enjoyment ratings (the 10 students in virtual condition and 14 students in mixed-reality condition that answered "I liked it very much"), we still find a significant, favorable effect of mixed-reality ($p = 0.001$).

Fig. 12. Children in mixed-reality condition (above) used more shape-relevant gestures while explaining their predictions than those in virtual condition (below).

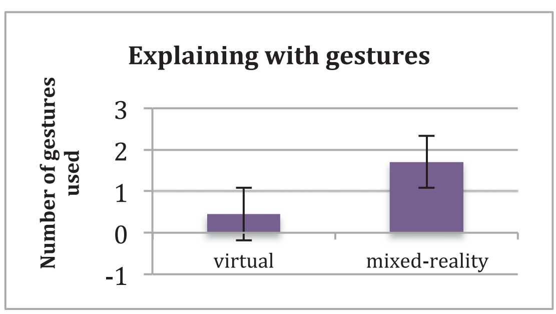

Fig. 13. Average number of meaningful gestures used to explain predictions.

4.4. Gestures as Signs of Mental Visualizations

Based on Alibali's theory that gestures can be signs of mental visualizations and embodied language [Hostetter and Alibali 2008], we used a measure of children's gestures as a proxy for mental visualizations. Pine et al. have also suggested that gestures can be an indicator of readiness to learn and cognitive gains [Pine et al. 2004]. While analyzing the videos, we noticed that the children in the mixed-reality condition used more gestures to explain their predictions. Specifically, these gestures often seemed to be encoding a tower's structure. 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 12). In the virtual condition, students mostly explained their predictions by pointing at the screen rather than using gestures that mirrored properties of the towers. The ANOVA analysis of our video data for students explaining their predictions revealed that students in the mixed-reality condition used significantly more gestures than those in the virtual condition ($p = 0.001$, $d = 0.72$). We counted only the gestures referring to the tower's structures, and did not count pointing gestures in our analysis (Figure 13). For the statistical analysis,

Table I. Percentage of Students Using Meaningful Gestures while Explaining Their Predictions and Learning

	No gesturing	Gesturing
Learning	0.34	0.23
No learning	0.38	0.05

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' 3D 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.

There was no significant correlation between gestures and learning. There are many students who do not gesture, and some of these students do learn (i.e., improve their scores between the paper pre-test and post-test; see the "No gesturing" column in Table I). However, there are very few students who gesture but do not learn, as shown in the "Gesturing" column in Table I. 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 students do gesture, it is a sign of their mental visualizations that is associated with better learning.

4.5. Qualitative Evidence

Qualitative anecdotes illustrate the students' enjoyment and engagement. Many children commented after the game that they liked the earthquake 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. Many children said that they would like to test their own towers on the earthquake table, suggesting that open-ended experimentation may lead to even more enjoyment. Further, although the pairs condition was designed to be collaborative, some students indicated that they would enjoy competing to build a tower that stayed up longest.

All of the a-ha moments occurred in the mixed-reality condition. Most of these a-ha moments happened 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!" We 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 (since we did not observe any a-ha moments in the virtual condition). One child also commented that she would prefer seeing the towers fall in real life rather than having a video or the computer say what happens.

The interface appeared intuitive for the children. They did not have questions about how to use it and did not demonstrate frustration with the interface design. Further, the children in the mixed-reality condition did not question how the gorilla character knew which tower had fallen.

5. EXPERIMENT 2: MIXED-REALITY AND PHYSICAL TRIGGER

The learning benefits that students derive from physical objects can be separated into three categories: benefits from physical observation; benefits from physical manipulation, in a manner that directly connects to the learning goals; and benefits from physical controls/triggers that are not directly connected to the learning goals. Experiment 1 showed that students do indeed benefit from observation: there were greater learning gains when the 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 replicates the finding that observing physical objects is beneficial for learning, and also explores the possible benefits of physical triggers that are not directly related to the target content. One pathway for such a benefit could be through enjoyment: children may enjoy interacting with direct physical triggers more than interacting with a mouse, and this enjoyment may lead to better learning.

5.1. Experimental Method

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 trigger [Yannier et al. 2015]. The goal of Experiment 2 was to test if adding a physical trigger like pressing a switch or shaking a tablet could increase learning and whether it might do so through increased enjoyment. Experiment 1 found no significant differences for learning or enjoyment between the solo and pair conditions but children and parents seemed to show a preference for playing in pairs. Thus, in Experiment 2 all participants played in pairs. Experiment 2 used the same tests and surveys as Experiment 1 to measure enjoyment and learning gains. The physical trigger in the mixed-reality game consisted of a physical switch that the children pressed to shake the table. The screen-based version was implemented on a tablet, which children physically shook to shake the virtual table. Similarly, we chose to have children press a physical switch to shake the table in the mixed-reality condition, based on the observation that in pilot tests, children seemed excited about pressing the physical switch to shake the table, suggesting that a physical trigger may lead to greater enjoyment.

5.1.1. Experimental Conditions. We 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 trigger (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 trigger (shaking the tablet as input). In each condition, students played in pairs. We discuss each in more detail below.

Mixed-reality version of EarthShake with mouse control: This condition was equivalent to the mixed-reality and 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 trigger to shake the earthquake table.

Mixed-reality version of EarthShake with physical trigger: This condition is identical to the mixed-reality version of EarthShake with mouse control, except that to shake the table, the children used a physical switch, connected wirelessly to the earthquake table (Figure 14).

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

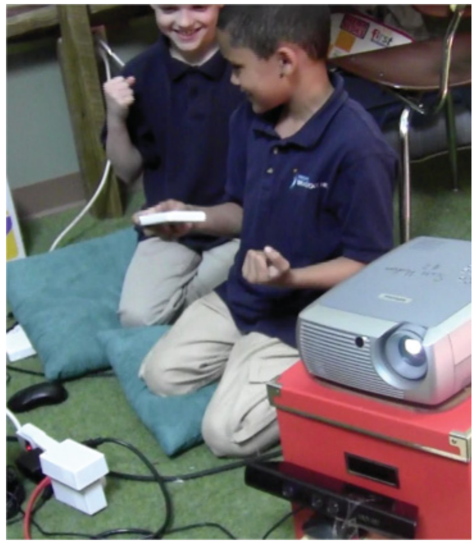

Fig. 14. Students in the mixed-reality, physical trigger version of EarthShake using a physical switch to shake the table.

Fig. 15. Students in the Virtual, physical trigger condition. Shaking the tablet makes the on-screen table shake.

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.

Screen-only version of EarthShake with mouse control: This condition was the same as the virtual and 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.

Screen-only tablet version of EarthShake with physical trigger: 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 trigger: for each of the 10 trials, children shook the tablet with their hands to activate the video of the towers falling (Figure 15). In this condition, the partners were asked to sit on the floor next to each other in a way that would allow both of

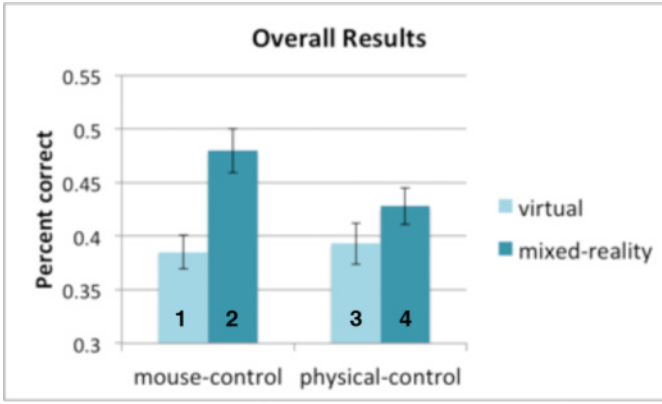

Fig. 16. Overall post-test results.

them to see the screen of the tablet. They took turns shaking the tablet and clicking on the selection choices.

5.1.2. Participants. 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. Pairs were selected by the teachers.

5.1.3. Procedure. The same procedure was used as in Experiment 1.

5.1.4. Measures. The same measures were used as in Experiment 1.

5.2. Results of Experiment 2

Paper pre- and post-tests and tower pre- and post-tests were analyzed to measure the learning gains from the experiment and investigate the learning effects of observing physical phenomena and using a physical trigger. Surveys were analyzed as a measure of enjoyment.

5.2.1. Paper-Test Results. A 2-way ANOVA with overall pre-test score as the outcome variable found no significant differences between the conditions at pre-test (F's < 0.46 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 trigger) and media type (mixed-reality or screen-only), with total pre-test score as a covariate and total post-test score as the outcome variable. 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 post-tests (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: the mixed-reality game improved learning almost five times more than 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 trigger conditions (Figure 16). This result indicates that, for young children, physical observation can improve learning, whereas a simple physical trigger is unlikely to.

It may be argued that the physical controls used in the experiment (shaking the tablet and pressing the switch to shake the tablet) are different forms of interaction

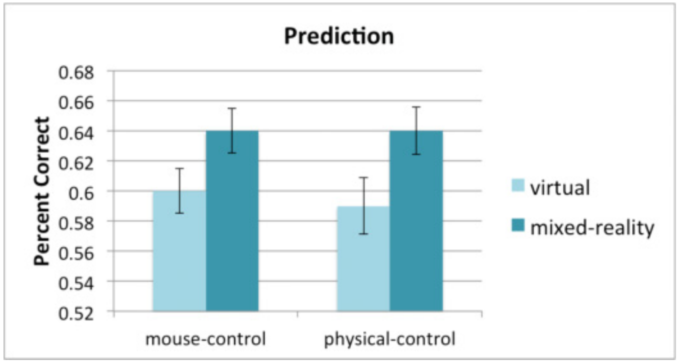

Fig. 17. Post-test scores for prediction items.

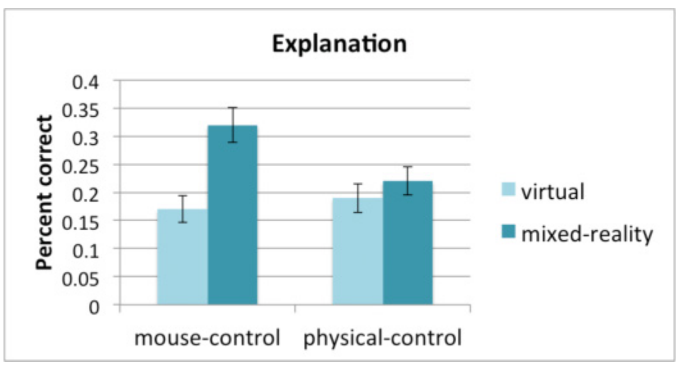

Fig. 18. Post-test scores for explanation items.

and should be analyzed on their own. Considering the conditions separately, we can see that the mouse-control mixed-reality condition (#2 in Figure 16) is significantly better than a typical virtual (#1) ($p < 0.05$), whereas the virtual, physical-control condition (#3) is not. Within the virtual conditions, there were no significant differences between the mouse control and the physical trigger (#1 and #3), and within the mixed-reality conditions, there is a non-significant trend in favor of mouse control. Thus, for students' learning, the effect of observing physical phenomena was more powerful than the effect of using a simple physical trigger such as shaking a tablet.

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 1% for virtual ($F(1,91) = 4.2$, $p < 0.05$, $d = 0.41$). The average post-test scores for the mixed-reality and virtual conditions were 64% and 60%, respectively (Figure 17). There was no significant effect of control type and also 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 18; 27% vs. 18% for post-test 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. Although the interaction between control type and media type

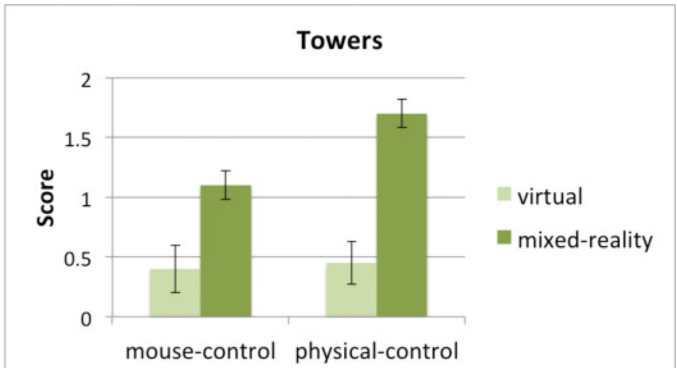

Fig. 19. Tower scores.

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.

Furthermore, we also analyzed the subset of participants who each gave their game the maximum likability rating (14 in the mixed-reality condition and 10 in the virtual), and still found a significant effect of media type on learning ($p = 0.001$), replicating the result from Experiment 1. This shows that even for those participants who were enjoying the game a lot, there was still a significant effect of physical observation on their learning.

5.2.2. Tower-Test Results. 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 19).

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 control types. This result is interesting as it shows that the benefits of physical observation transfer to a constructive problem-solving task involving physical interaction.

5.3. Enjoyment

Enjoyment was measured with the same survey as in Experiment 1 (Figure 20). 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 significant interaction of media type and control type. Though the interaction was not significant, we do observe a trend among students in the mixed-reality conditions in the direction of greater enjoyment for the physical-control students than the mouse-control students. On the other hand, there is no such trend for the virtual conditions.

Although it seems reasonable that increased enjoyment may improve learning, we did not find evidence supporting the idea that increased enjoyment produces increased learning. There was no correlation between enjoyment and pre–post learning gains ($r = 0.16$).

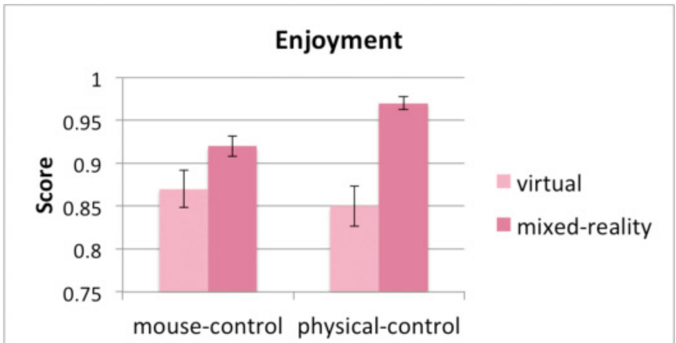

Fig. 20. Enjoyment scores based on the survey.

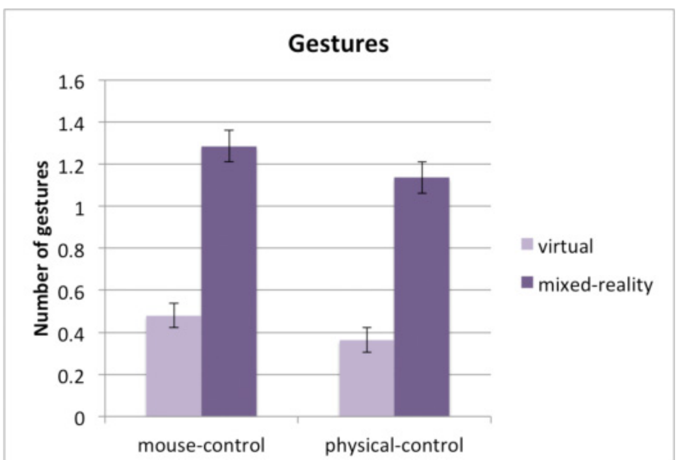

Fig. 21. Meaningful gestures children used during their explanations.

5.4. Gestures and Signs of Mental Visualizations

We examined the gestures students made while making their predictions and explaining the results (why one of the towers fell first). To measure indications of embodied cognition and mental simulations, we coded those gestures with the same scheme as in Experiment 1. Students in the mixed-reality condition used a greater number of meaningful gestures than those in the virtual condition. An ANOVA shows that this difference is significant ($F(1,92) = 11.55$, $p = 0.001$, $d = 0.72$) (Figure 21). This result is consistent with the gesture results from the first experiment.

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 and Alibali 2008]. The significant correlation in our 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.

5.5. Qualitative Evidence

The qualitative evidence in this experiment was similar to that in Experiment 1. Again, children seemed to be very 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 said 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 said that she thought this was like the next version of smart boards.

For some children, EarthShake engaged their imaginations as well. 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. Seeing real towers may trigger children's imagination, facilitating embodiment of their stories and help them make connections with familiar objects. All of these pathways may facilitate learning [Manches and Price 2011].

6. GENERAL DISCUSSION

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 post-test 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, these results also suggest that some of the learning benefits derived from students' manipulation of physical objects come from simply observing those objects. Experiment 2 replicated this intriguing result and also showed that a simple physical trigger, such as shaking a tablet or pressing a switch, does not have a significant effect on learning or enjoyment. Although students may benefit from manipulating physical objects in ways that directly relate to the content they are learning, the results from Experiment 2 suggest that simple physical triggers that are not inherently related to the content will not lead to additional benefits.

Furthermore, our results revealed that the learning benefits for prediction and explanation transfer to a construction task as well. The towers of the mixed-reality conditions improved significantly more than those in the virtual conditions. Thus, the mixed-reality conditions not only fostered better learning of the balance principles, but also better application of those principles in a hands-on, constructive problem-solving task.

To our knowledge, this paper 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 a computer game. In most of the prior studies on tangible interfaces, users had both observed and manipulated the tangible objects. In contrast, this study disentangles those variables, and establishes a learning benefit for simply observing a physical phenomenon. 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? We explored three theoretical explanations for why observing physical phenomenon may produce more science learning: that the presence of physical objects is inherently more enjoyable, that it facilitates embodied cognition and 3D mental visualizations, and that it enhances collaboration.

Our data do 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. Students in the mixed-reality condition both qualitatively showed more enjoyment and rated their enjoyment higher on the quantitative survey. However, although our data support the theory that interaction with physical objects is more enjoyable than interaction with virtual ones, increased enjoyment does not seem sufficient to explain the large learning differences we observed. We found that among the subset of participants who gave their game the maximum likability rating (14 in the mixed-reality condition and 10 in the virtual), there was still a significant effect of media type on learning ($p = 0.001$). This result was replicated for both experiments. We acknowledge that Likert scale measurements are limited and may not be the most reliable measure for enjoyment. Even though we acknowledge that we cannot completely discard the enjoyment explanation based on our survey results, our data do not seem to provide evidence to support the enjoyment explanation.

Our gesture data provide some support for the explanation that observing physical phenomena in the real world supports embodied cognition, and helps children perceive, mentally visualize, and ultimately remember concepts better. Children in the mixed-reality conditions more often explained their predictions using meaningful gestures that indicated 3D mental visualizations and 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 2013]; and (2) the spontaneous gestures children produce when explaining a task are a sign of their mental visualizations and predict how much they will learn from that task [Hostetter and Alibali 2008].

An alternative explanation for the enhanced learning in the mixed-reality condition is that those students were given more information: they saw both the physical towers and virtual representations of those towers. Perhaps the mental processing required to map the physical towers and their corresponding 2D screen displays caused the deeper learning in the mixed-reality condition. However, it is not clear if such processing is germane or extraneous to learning the principles of balance and stability. If the processing is extraneous, it may lead to split attention and have negative consequences for learning [Ayres and Sweller 2014].

In sum, our current evidence is most supportive of the theory that physical observation in the real world facilitates 3D mental visualizations and enhances retrieval and reasoning. We do not have any evidence supporting the hypothesis that the results are merely a consequence of increased enjoyment. Our data also suggest that simple physical triggers such as shaking a tablet do not improve learning or enjoyment. However, in this experiment we do not investigate the effect of hands-on manipulation of physical objects. Piaget's theory stresses the potential role of physical manipulation in developing understanding. Physical manipulation could also aid collaboration through the affordance and awareness of sharing objects [Yuill and Rogers 2012]. In future work, we plan to integrate more hands-on activities into EarthShake and investigate the role of hands-on exploration vs. guided discovery on learning and enjoyment.

7. DESIGN IMPLICATIONS

Our studies show the impact of physical observation coupled with immediate interactive feedback on children's science learning and understanding of physical phenomena.

The design iterations we 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. This feedback was designed to scaffold students' construction of their own explanations and foster understanding of early physics principles without being told directly how they applied. In the pilot studies where we used the earthquake table on its own, without the projected game, it appeared children had less success in learning the physics principles. In contrast, when they saw the self-explanation menu in the background while also seeing the physical towers in the foreground, they were able to recognize the relevant principle (such as having more weight on top than bottom) even if they had not predicted it beforehand. This suggests that the self-explanation menu synchronized with the physical world was a critical component of the game and facilitated learning. This self-explanation menu can be generalized to other mixed-reality environments.

We realized that children enjoyed 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 said explicitly that he would like it better if the game had more buildings. Thus, incorporating more hands-on activities in the central game mechanic (and addressing the associated technical challenges) may yield further benefits. In future versions, we hope to integrate more building activities into the game.

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 we made so that they would hear all the answers without skipping through them). One of the children complained, “I don't want the gorilla to speak so much!”

We observed that some of the children in the pair conditions had a hard time sharing the mouse with their partner. Interactions for pairs may be improved with a more tangible approach for the selection of menu items (e.g., allowing students to select items by pointing or with a touch screen).

EarthShake, which consists of a physical experimental setup, depth-camera sensing, and a projected game/activity can be extended for many different content areas outside the domain of balance and stability (e.g., projectile motion, density, human body, planetary systems, math, and so on). As future work, we plan to create new mixed-reality games in different content areas, which use a similar setup (the Kinect camera, projector, and physical experimentation). Eventually, we aim to create a mixed-reality platform bridging the advantages of virtual and physical worlds via inexpensive depth-camera sensing, which can be reused for different content areas in education.

7.1. Why is the Technology Critical?

It is important to emphasize that there are important elements of the technology in our 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 our 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 to help him/her 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. As described in

Section 4.5, seeing the explanation menu prompted a-ha moments 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!" We 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.

The feedback provided in the game is critical. Games in this domain without such support lead to little learning (e.g., Rumbleblocks [Christel et al. 2012]). Further, there is much evidence that children in general learn better with guidance, in particular with feedback and self-explanation [Aleven and Koedinger 2000]. Without scaffolding and support, students often miss the point of the learning activity [Puchner et al. 2001]. Particularly, in science, the phenomenon of "confirmation bias" [Nickerson 1998] suggests that children are likely to see their predictions as confirmed even when they are not, and explicit indication otherwise can reduce this tendency. The EarthShake system uses depth-camera sensing to provide personalized immediate feedback on real world phenomena, allowing children to discover new principles with some support and scaffolding.

8. CONCLUSION

This paper examines EarthShake, a mixed-reality game that uses depth-camera sensing to bring together physical and virtual worlds. It also incorporates guided learning and self-explanation prompts. Results from our two 2×2 experiments show that the mixed-reality version of EarthShake helped children learn more than otherwise-equivalent flat-screen versions (laptop or tablet) of the same game. Children interacting with the mixed-reality game reported greater enjoyment and used more meaningful gestures to explain their predictions. These results indicate that the observation of physical phenomena in an interactive game can foster effective learning and inferences. Thus, tangible interfaces and mixed-reality learning environments that provide such affordances have great potential not only to increase enjoyment but also to improve children's learning, especially in science. Our results are also of theoretical interest: first, analysis of students' gestures supports explanation of learning benefits through mental visualization. Second, our tightly-controlled design suggests that in the context of camera-synchronized interaction, children gain some of the benefits of embodied cognition through mere observation of physical phenomena.

ACKNOWLEDGMENTS

We would like to thank Cathy Chase, Vincent Aleven, Kelly Rivers, Rony Patel, Chris MacLellan, and the ENGAGE team for their kind help with this work. We would also like to thank the participating schools, teachers, and students.

REFERENCES

Dor Abrahamson, Rosa G. Lee, Andrea G. Negrete, and José F. Gutiérrez. 2014. Coordinating visualizations of polysemous action: Values added for grounding proportion. *ZDM* 46, 1 (2014), 79–93.

Vincent A. W. M. M. Aleven and Kenneth R. Koedinger. 2000. The need for tutorial dialog to support self-explanation. In *Paper Presented at the AAAI Fall Symposium on Building Dialogue Systems for Tutorial Applications*. 65–73.

Alissa N. Antle. 2013. Research opportunities: Embodied child–computer interaction. *Int. J. Child-Comput. Interact.* 1, 1 (2013), 30–36. DOI: http://dx.doi.org/10.1016/j.ijcci.2012.08.001

Alissa N. Antle, Milena Droumeva, and Daniel Ha. 2009. Hands on what?: Comparing children's mouse-based and tangible-based interaction. In *Proceedings of the 8th International Conference on Interaction Design and Children*. 80–88.

Paul Ayres and John Sweller. 2014. The split-attention principle in multimedia learning. *Cambridge Handb. Multimed. Learn.* 2 (2014).

Margarita Azmitia and Kevin Crowley. 2001. The rhythms of scientific thinking: A study of collaboration in an earthquake microworld. *Des. Sci. Implic. from everyday, classroom, Prof. settings* (2001), 51–81.

Maribeth Back, Jonathan Cohen, Rich Gold, Steve Harrison, and Scott Minneman. 2001. Listen reader: An electronically augmented paper-based book. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 23–29.

Saskia Bakker, Elise van den Hoven, and Alissa N. Antle. 2011. MoSo tangibles: Evaluating embodied learning. In *Proceedings of the 5th International Conference on Tangible, Embedded, and Embodied Interaction.* 85–92.

Rachel Barr. 2010. Transfer of learning between 2D and 3D sources during infancy: Informing theory and practice. *Dev. Rev.* 30, 2 (2010), 128–154.

Hrvoje Benko, Ricardo Jota, and Andrew Wilson. 2012. MirageTable: freehand interaction on a projected augmented reality tabletop. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 199–208.

Robert M. Carini, George D. Kuh, and Stephen P. Klein. 2006. Student engagement and student learning: Testing the linkages. *Res. High. Educ.* 47, 1 (2006), 1–32.

Catherine C. Chase, Jonathan T. Shemwell, and Daniel L. Schwartz. 2010. Explaining across contrasting cases for deep understanding in science: An example using interactive simulations. In *Proceedings of the 9th International Conference of the Learning Sciences*, Vol. 1. 153–160.

Michael G. Christel, Scott M. Stevens, Bryan S. Maher, Sean Brice, Matt Champer, Luke Jayapalan, Qiaosi Chen, Jing Jin, Daniel Hausmann, Nora Bastida, and Xun Zhang. 2012. RumbleBlocks: Teaching science concepts to young children through a Unity game. In *Proceedings of the 17th International Conference on Computer Games (CGAMES'17)*. 162–166.

Albert T. Corbett and John R. Anderson. 2001. Locus of feedback control in computer-based tutoring: Impact on learning rate, achievement and attitudes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 245–252.

Brent Davis. 2015. *Spatial Reasoning in the Early Years: Principles, Assertions, and Speculations*, Routledge.

Johannes Engelkamp and Hubert D. Zimmer. 1989. Memory for action events: A new field of research. *Psychol. Res.* 51, 4 (1989), 153–157.

Taciana Pontual Falcão and Sara Price. 2009. What have you done! the role of'interference'in tangible environments for supporting collaborative learning. In *Proceedings of the 9th International Conference on Computer Supported Collaborative Learning*, Vol. 1. 325–334.

George W. Fitzmaurice, Hiroshi Ishii, and William A. S. Buxton. 1995. Bricks: Laying the foundations for graspable user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 442–449.

Arthur M. Glenberg. 1997. What memory is for: Creating meaning in the service of action. *Behav. Brain Sci.* 20, 01 (1997), 41–50.

Anuradha A. Gokhale. 1995. Collaborative learning enhances critical thinking. *Journal of Technology Education.*

Ankit Gupta, Dieter Fox, Brian Curless, and Michael Cohen. 2012. DuploTrack: A real-time system for authoring and guiding duplo block assembly. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*. 389–402.

Harlene Hayne, Jane Herbert, and Gabrielle Simcock. 2003. Imitation from television by 24-and 30-month-olds. *Dev. Sci.* 6, 3 (2003), 254–261.

Philip H. Henning. 2004. Everyday cognition and situated learning. *Handb. Res. Educ. Commun. Technol.* (2004), 143–168.

Eva Hornecker and Jacob Buur. 2006. Getting a grip on tangible interaction: A framework on physical space and social interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 437–446.

Autumn B. Hostetter and Martha W. Alibali. 2008. Visible embodiment: Gestures as simulated action. *Psychon. Bull. Rev.* 15, 3 (2008), 495–514.

David Klahr, Lara M. Triona, and Cameron Williams. 2007. Hands on what? The relative effectiveness of physical versus virtual materials in an engineering design project by middle school children. *J. Res. Sci. Teach.* 44, 1 (2007), 183–203.

Kenneth R. Koedinger, Elizabeth A. McLaughlin, and Neil T. Heffernan. 2010. A quasi-experimental evaluation of an on-line formative assessment and tutoring system. *J. Educ. Comput. Res.* 43, 4 (2010), 489–510.

Harlan Lebo. 2007. *The 2007 Digital Future Report: Surveying the Digital Future, Year Six*. Center for the Digital Future, USC Annenberg School.

Mark W. Lipsey, Kelly Puzio, Cathy Yun, Michael A. Hebert, Kasia Steinka-Fry, Mikel W. Cole, Megan Roberts, Karen S. Anthony, and Matthew D. Busick. 2012. *Translating the Statistical Representation of the Effects of Education Interventions into More Readily Interpretable Forms*. National Center for Special Education Research.

Andrew Manches. 2011. Digital manipulatives: Tools to transform early learning experiences. *Int. J. Technol. Enhanc. Learn.* 3, 6 (2011), 608–626.

Andrew Manches and Sara Price. 2011. Designing learning representations around physical manipulation: hands and objects. In *Proceedings of the 10th International Conference on Interaction Design and Children*. 81–89.

Paul Marshall, Peter C. H. Cheng, and Rosemary Luckin. 2010. Tangibles in the balance: A discovery learning task with physical or graphical materials. In *Proceedings of the Fourth International Conference on Tangible, Embedded, and Embodied Interaction*. 153–160.

Taylor Martin and Daniel L. Schwartz. 2005. Physically distributed learning: Adapting and reinterpreting physical environments in the development of fraction concepts. *Cogn. Sci.* 29, 4 (2005), 587–625.

Maria Montessori. 1964. *The Montessori Method*. Transaction publishers.

Raymond S. Nickerson. 1998. Confirmation bias: A ubiquitous phenomenon in many guises. *Rev. Gen. Psychol.* 2, 2 (1998), 175.

Claire O'Malley and Danae Stanton Fraser. 2004. *Literature Review in Learning with Tangible Technologies*. A NESTA Futurelab Research report - report 12.

Georgios Olympiou and Zacharias C. Zacharia. 2012. Blending physical and virtual manipulatives: An effort to improve students' conceptual understanding through science laboratory experimentation. *Sci. Educ.* 96, 1 (2012), 21–47. DOI:http://dx.doi.org/10.1002/sce.20463

Karen J. Pine, Nicola Lufkin, and David Messer. 2004. More gestures than answers: Children learning about balance. *Dev. Psychol.* 40, 6 (2004), 1059.

Sara Price, George Roussos, Taciana Pontual Falcão, and Jennifer G. Sheridan. 2009. Technology and embodiment: Relationships and implications for knowledge, creativity and communication. *Beyond Curr. Horiz.* (2009), 85–92.

Laurel Puchner, Robyn Rapoport, and Suzanne Gaskins. 2001. Learning in children's museums: Is it really happening? *Curator Mus. J.* 44, 3 (2001), 237–259.

Helen Quinn, Heidi Schweingruber, Thomas Keller, and others. 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. National Academies Press.

Mitchel Resnick, et al. 1998a. Digital manipulatives. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'98)*. 281–287. DOI:http://dx.doi.org/10.1145/274644.274684

Mitchel Resnick, Fred Martin, Robert Berg, Rick Borovoy, Vanessa Colella, Kwin Kramer, and Brian Silverman. 1998b. Digital manipulatives: New toys to think with. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'98)*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 281–287. DOI:http://dx.doi.org/10.1145/274644.274684

John J. Rieser, Anne E. Garing, and Michael F. Young. 1994. Imagery, action, and young children's spatial orientation: It's not being there that counts, it's what one has in mind. *Child Dev.* 65, 5 (1994), 1262–1278.

Keith Roe and Daniel Mujis. 1998. Children and computer games – A profile of the heavy user. *Eur. J. Commun.* 2, 13 (1998), 181–200.

Yvonne Rogers, Mike Scaife, Silvia Gabrielli, Hilary Smith, and Eric Harris. 2002. A conceptual framework for mixed reality environments: Designing novel learning activities for young children. *Presence* 11, 6 (2002), 677–686. DOI:http://dx.doi.org/10.1162/105474602321050776

Kimiko Ryokai and Justine Cassell. 1999. StoryMat: A play space for collaborative storytelling. In *Extended Abstracts on Human Factors in Computing Systems (CHI EA'99)*. ACM, New York, NY, USA, 272–273. DOI:http://dx.doi.org/10.1145/632716.632883

Bertrand Schneider, Patrick Jermann, Guillaume Zufferey, and Pierre Dillenbourg. 2011. Benefits of a tangible interface for collaborative learning and interaction. *IEEE Trans. Learn. Technol.* 4, 3 (2011), 222–232. DOI:http://dx.doi.org/10.1109/TLT.2010.36

Tia Shelley, Leilah Lyons, Moira Zellner, and Emily Minor. 2011. Evaluating the embodiment benefits of a paper-based tui for educational simulations. In *Extended Abstracts on Human Factors in Computing Systems (CHI EA'11)*. ACM, New York, NY, USA, 1375–1380. DOI:http://dx.doi.org/10.1145/1979742.1979777

D. Stanton, C. O'Malley, K. Huing, M. Fraser, and S. Benford. 2003. Situating historical events through mixed reality. In Barbara Wasson, Sten Ludvigsen, and Ulrich Hoppe (Eds.). *Designing for Change in Networked Learning Environments: Proceedings of the International Conference on Computer Support for Collaborative Learning*. Springer Netherlands, Dordrecht, 293–302. DOI:http://dx.doi.org/10.1007/978-94-017-0195-2_37

Tiffany Tseng, Coram Bryant, and Paulo Blikstein. 2011. Collaboration through documentation: automated capturing of tangible constructions to support engineering design. In *Proceedings of the 10th International Conference on Interaction Design and Children (IDC'11)*. ACM, New York, NY, USA, 118–126. DOI:http://dx.doi.org/10.1145/1999030.1999044

B. Ullmer and H. Ishii. 2000. Emerging frameworks for tangible user interfaces. *IBM Syst. J.* 39, 3–4 (Jul. 2000), 915–931. DOI:http://dx.doi.org/10.1147/sj.393.0915

David H. Uttal and Katherine O' Doherty. 2008. Comprehending and learning from 'visualizations': A developmental perspective. In John K. Gilbert, Miriam Reiner, and Mary Nakhleh (Eds.). *Visualization: Theory and Practice in Science Education*. Dordrecht: Springer Netherlands, 53–72. DOI:http://dx.doi.org/10.1007/978-1-4020-5267-5_3

Consuelo Valdes, Michelle Ferreirae, Taili Feng, Heidi Wang, Kelsey Tempel, Sirui Liu, and Orit Shaer. 2012. A collaborative environment for engaging novices in scientific inquiry. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS'12)*. ACM, New York, NY, USA, 109–118. DOI:http://dx.doi.org/10.1145/2396636.2396654

Lev Vygotsky. 1978. Interaction between learning and development. *Read. Dev. Child.* 23, 3 (1978), 34–41.

Erin Walker and Winslow Burleson. 2012. Using need validation to design an intelligent tangible learning environment. *Proc. ACM Annual Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA'12)*, 2123. DOI:http://dx.doi.org/10.1145/2212776.2223763

Nesra Yannier, Cagatay Basdogan, Serdar Tasiran, and Omer L. Sen. 2008. Using haptics to convey cause-and-effect relations in climate visualization. *IEEE Trans. Haptics* 1, 2 (2008), 130–141. DOI:http://dx.doi.org/10.1109/TOH.2008.16

Nesra Yannier, Kenneth R. Koedinger, and Scott E. Hudson. 2015. 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 (CHI'15)*. ACM, New York, NY, USA, 1045–1054. DOI:http://dx.doi.org/10.1145/2702123.2702397

Nesra Yannier, Kenneth R. Koedinger, and Scott E. Hudson. 2013. Tangible collaborative learning with a mixed-reality game: EarthShake. In H. Chad Lane, Kalina Yacef, Jack Mostow, and Philip Pavlik, (Eds.). *Artificial Intelligence in Education: Proceedings of the 16th International Conference (AIED'13) Memphis, TN, USA, July 9–13, 2013*. Springer, Berlin, 131–140. DOI:http://dx.doi.org/10.1007/978-3-642-39112-5_14

Nicola Yuill and Yvonne Rogers. 2012. Mechanisms for collaboration: A design and evaluation framework for multi-user interfaces. *ACM Trans. Comput. Interact.* 19, 1 (May 2012), 1:1–1:25. DOI:http://dx.doi.org/10.1145/2147783.2147784

Received May 2015; revised April 2016; accepted May 2016