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


**Nesra Yannier**

Human-Computer Interaction Institute
School of Computer Science
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213



CMU-HCII-16-103

August, 2016


**Thesis Committee:**

Kenneth R. Koedinger, Carnegie Mellon University (Co-Chair)
Scott E. Hudson, Carnegie Mellon University (Co-Chair)
Jessica Hammer, Carnegie Mellon University
Kevin Crowley, University of Pittsburgh

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy




This work was supported partly by contract ONR N00014-12-C-0284 (awarded by DARPA), partly by LearnLab (NSF grant SBE-0836012), and by Mister Rogers Scholarship, Sprout's EdTech Refinery Grant and an Innovation Fellowship by Center for Innovation and Entrepreneurship.

# Keywords

Tangible Interfaces, Mixed-Reality Learning, enjoyment, classroom experiments, user-centered design, controlled experiments, contrasting cases, guided discovery, exploration, interactive feedback, depth-camera sensing, games, scientific inquiry

# Abstract

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

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

My thesis work further investigates what features of a mixed-reality system yield better learning and enjoyment, especially in the context of limited experimental results from other mixed-reality learning research. Most mixed-reality environments, including tangible interfaces, currently emphasize open-ended exploration and problem solving, and are claimed to be most effective when used in a discovery-learning mode with minimal guidance. I investigated how critical to learning and enjoyment interactive guidance and feedback is (e.g. predict/observe/explain prompting structure with interactive feedback), in the context of EarthShake. In a third experiment, I compared the learning and enjoyment outcomes of

children interacting with a version of EarthShake that supports guided-discovery, another version that supports exploration in discovery-learning mode, and a version that is a combination of both guided-discovery and exploration. The results of the experiment reveals that Guided-discovery and Combined conditions where children are exposed to the guided discovery activities with the predict-observe-explain cycle with interactive feedback yield better explanation and reasoning. Thus, having guided-discovery in a mixed-reality environment helps with formulating explanation theories in children's minds. However, the results also suggest that, children are able to activate explanatory theory in action better when the guided discovery activities are combined with exploratory activities in the mixed-reality system. Adding exploration to guided-discovery activities, not only fosters better learning of the balance/physics principles, but also better application of those principles in a hands-on, constructive problem-solving task.

My dissertation contributes to the literatures on the effects of physical observation and mixed-reality interaction on students' science learning outcomes in learning technologies. Specifically, I have shown that a mixed-reality system (i.e., combining physical and virtual environments) can lead to superior learning and enjoyment outcomes than screen-only alternatives, based on different measures. My work also contributes to the literature of exploration and guided-discovery learning, by demonstrating that having guided discovery activities in a mixed-reality setting can improve children's fundamental principle learning by helping them formulate explanations. It also shows that combining an engineering approach with scientific thinking practice (by combining exploration and guided-discovery activities) can lead to better engineering outcomes such as transferring to constructive hands-on activities in the real world. Lastly, my work aims to make a contribution from the design perspective by creating a new mixed-reality educational system that bridges physical and virtual environments to improve children's learning and enjoyment in a collaborative way, fostering productive dialogue and scientific curiosity in museum and school settings, through iterative evolving design methodology to ensure effective learning and enjoyment outcomes in these settings.

# Table of Contents:

# Chapter 1. Introduction

Today even very young children are being drawn into the compelling world of screen-based technologies, such as tablets or computer games. Screen-based technologies can help children learn by providing immediate targeted feedback [Corbett and Anderson 2001]. However, as screen-based technologies are becoming so appealing for children, it is worth asking whether real world interaction is still needed to enhance learning and enjoyment. Are today's children actually missing out on opportunities to develop understanding more readily in their physical environment by being immersed in these flat-screen technologies? In fact, some have argued that these flat-screen technologies can have negative effects on children [Roe and Mujis 1998; Lebo 2007] [Lave 1991][Turkle 2015].

How can games for young children combine the distinct advantages of screen-based and physical learning environments? In addition to instructional support, screen-based games also have potential motivational benefits, such as compelling scenarios and engaging characters. On the other hand, most natural learning occurs in our physical 3D world and some argue that that is where learning is at its best [Henning 2004]. The physical world has the potential to help children play, discover, experiment and learn in their everyday world and do so in a way that also supports social interaction [Turkle 2015]. Indeed, many past technology efforts have encouraged children to play with physical objects such as building blocks and puzzles to learn a variety of skills [O'Malley 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 and mixed-reality environments may help students learn in more engaging and powerful ways than either approach alone.

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

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

## 1.1 Learning with Physical Objects

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

## 1.2 Mixed-Reality Environments and Tangible Interfaces for Learning

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

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

Unlike the mixed results for non-instrumented physical objects, research comparing tangible and virtual interactions generally shows a benefit for tangibles (mostly performance benefits rather than learning outcomes with pre/post tests). Children were more successful and faster at solving puzzles when using tangible puzzle pieces instead of comparable interactions with a mouse [Antle et al. 2009]. Bakker et al. designed and evaluated MoSo Tangibles: a set of interactive, physical artifacts with which children can manipulate the pitch, volume and tempo of ongoing tones, in order to structure their understanding of these abstract sound concepts. Their results indicate that MoSo provided children with a physical handle to reason about targeted abstract concepts (with qualitative interviews and video analysis) [Bakker et al. 2011]. Shelley et al. demonstrated problem solving and collaboration advantages for a paper-based tangible user interface for educational simulations over mouse interaction [Shelley et al. 2011]. Logistic apprentices demonstrated enhanced task performance, collaborative interactions, and sense of playfulness when using a tangible instead of multi-touch interface [Schneider et al. 2011]. In another study, students better remembered cause and effect relations in climate when they used a haptics-augmented environment where they could feel forces in addition to a virtual environment [Yannier et al. 2008]. In another study, Yannier et al. created and evaluated FeelSleeve, an interface that allows children to feel story events in their hands via haptic feedback while they are reading on a mobile device. Their results showed that story events accompanied by haptic feedback are better comprehended and appear to be more salient in memory [Yannier, Israr, et al. 2015].

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

### 1.3 Theoretical Background

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

First of all, experiencing a physical phenomenon may help people perceive and mentally visualize the target objects [Engelkamp and Zimmer 1989][Abrahamson et al. 2014], leading to better understanding of scientific principles underlying physical phenomena. This mental visualization may then facilitate connections with familiar objects, and result in improved memory for the concepts related to those objects. Physical observations may be more deeply processed so as to recognize key features that explain physical phenomena (e.g. that a higher center of mass leads to instability). This theory follows Antle's research on embodied child-computer interaction, suggesting that when children (and adults) learn or reason with abstract concepts, they utilize mental simulations based on concrete motor-perceptual experiences [Antle 2013]. Also Hostetter et al. have theorized that perceptual and motor simulations underlie embodied language and mental imagery, which are often revealed by spontaneous gestures that accompany speech [Hostetter 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 rerun, reactivating the same neural patterns [Antle 2013]. 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, observing phenomena in the real world in a mixed-reality environment may trigger mental simulations and affordances for action, facilitating retrieval from memory.

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

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

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

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

# Chapter 2. EarthShake

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

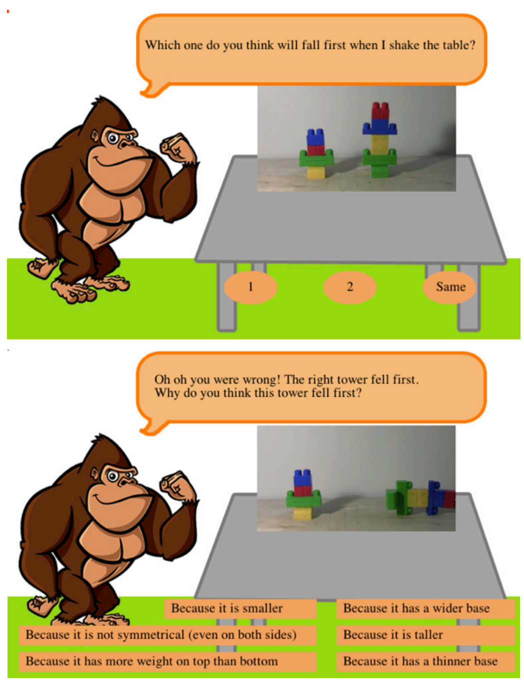



**Figure 1.** Virtual-only version of Earthshake, showing the predict/observe/explain cycle. The video of the physical towers shaking on the earthquake table is integrated into the game interface.

As shown in Fig 2, EarthShake consists of a multimodal interactive earthquake table, physical towers made of blocks, a Kinect depth camera and a display screen behind the table. It utilizes a predict/observe/explain cycle, where children are asked to make *predictions* about stability, *observe* outcomes of physical experiments, and *explain* those outcomes. The system detects which of the towers in the physical setup falls first when the user shakes the table and gives visual and audio feedback accordingly [Yannier et al. 2013]. Children are guided by pedagogical prompts that highlight whether or not a prediction was correct and that scaffold explanations of the actual outcome.

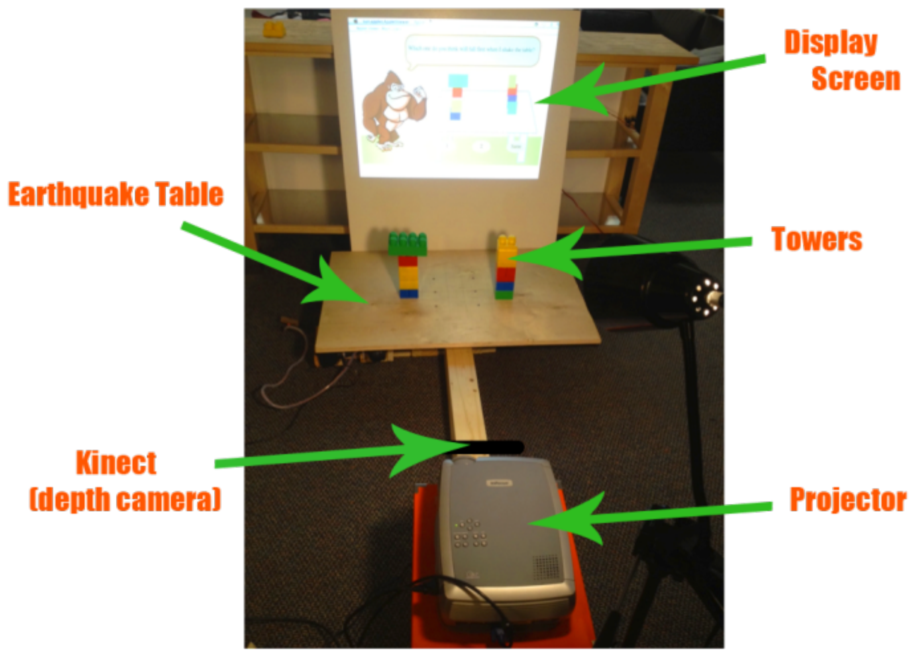



**Figure 2.** First version of the physical setup of EarthShake.

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

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

## 2.1 Scenario

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

When the table shakes, the Kinect camera and computer vision algorithm determine which tower fell. If the students' prediction was correct, the gorilla says: "Good job! Your hypothesis was right. Why do you think this tower fell first?" If they were wrong, he says: "Oh oh you were wrong! Why do you think this tower fell first?" To explain why that tower fell, the students choose one of six explanations projected on the screen, providing scaffolding. The menu, read aloud by the gorilla, consists of the following choices: "Because it is smaller", "Because it is taller", "Because it has more weight on top than bottom", “Because it has a wider base”, “Because it is not symmetrical”, “Because it has a thinner base”. (Figure 1) This scenario is repeated for different contrasting cases targeting height, wide base, symmetry, and center of mass principles (Figure 3). Note that while students observe the physical towers, they do not touch them.

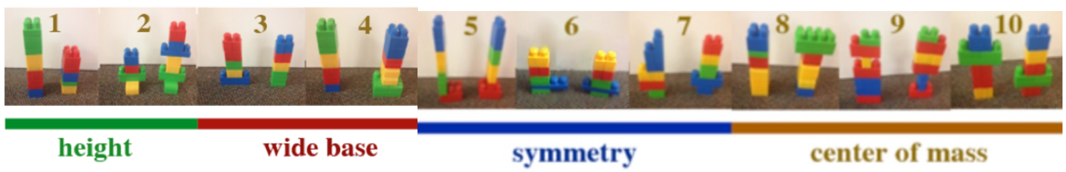



**Figure 3.** Contrasting cases used during the game.

### 2.2 Physical Setup and Vision Algorithm

The physical setup of EarthShake includes an earthquake table, physical towers placed on the table, a Kinect depth camera facing the towers, a projector, and a display screen with the computer game (Figure 3). The Kinect camera and our specialized computer vision algorithm detects when a tower falls, ensuring that EarthShake is in sync with what is happening in the real world. The projected computer game provides visual and audio feedback to the user (e.g., noting which tower the student predicted would fall and which actually fell) [Yannier et al. 2013]. Our technology and teaching method provides personalized interactive feedback to the users as they experiment and make discoveries in their physical environment [Yannier et al. 2016].

The earthquake table consists of a small motor, a switch/relay, a mechanism for converting from rotary to reciprocating linear motion and rails to support the reciprocating platform. When the switch or the relay is activated, it activates the motor, which then moves the platform back and forth.

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

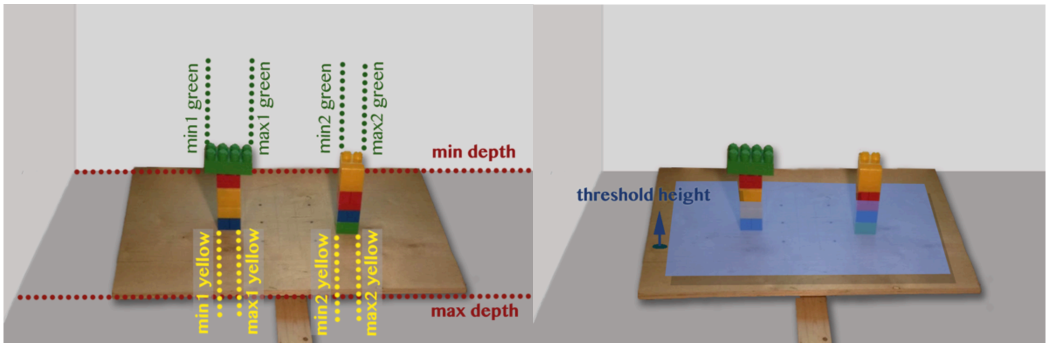


**Figure 4.** Specialized vision algorithm first calculates the minimum and maximum values for each color blob in each tower to determine where each tower stands (top). Then I use a threshold height to detect a fall, if all the color blobs in the tower are below this height (bottom).

In our first algorithm, relying on color information, caused problems in real world settings, as the lighting of the room affected the algorithm and caused inaccuracies. Therefore, in the second version of our algorithm, we decided to rely on more depth information than color information since the depth information does not change to according to lighting and is reliable in real world settings. In this version, again an image is extracted from the Kinect. We do filtering on the image to remove the background past a given depth. Then we store the blobs in the image in an array. A Moment of Inertia based metric (based on the formula below) is calculated for each object in the image and then the value is compared to the Tower Database, to determine which tower has been placed.

Moment of Inertia is a quantity expressing a body's tendency to resist angular acceleration. It is the sum of the products of the mass of each particle in the body with the square of its distance from the axis of rotation. It can be calculated with the following formula:

$$xMoment(blob) = \frac{\sum_{p\epsilon blob}(p.x - blob.minx)^2}{|blob|}$$

$$yMoment(blob) = \frac{\sum_{p\epsilon blob}(p.y - blob.miny)^2}{|blob|}$$

where this equation is summed up for the pixels in the array (i.e. pixelArray[i] which contains the x,y position of the ith pixel in the blob) and then normalized by dividing by the number of pixels in the blob. This process is repeated for each axis, which results in unique moment of inertia values for each object, which can then be used to distinguish between different objects (Figure 5).

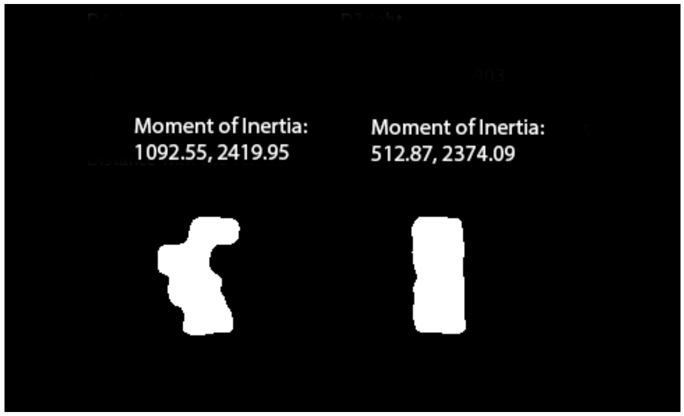



**Figure 5.** Physics based Moment of Inertia is calculated for each tower, which is a unique value that can help differentiate between different towers.

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 a specialized vision algorithm to overcome this challenge.

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

# Chapter 3. Pilot Study

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

Paper pre- and post-tests demonstrated large learning gains. On multiple-choice items asking students to predict which of two towers would fall first, 62% answered correctly at pretest, and 78% answered correctly at 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 pretest, while 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].

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

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

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

# Chapter 4.
# Experiment 1: Mixed-Reality and Collaboration

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

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

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

### 4.1 Procedure

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

### 4.2 Results of Experiment 1

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

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

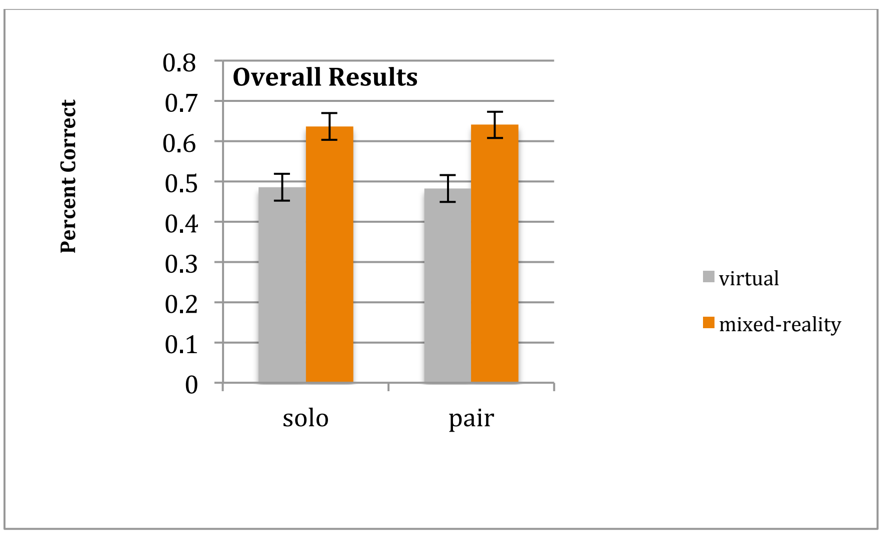



**Figure 6a.** Overall Post-test Learning Results

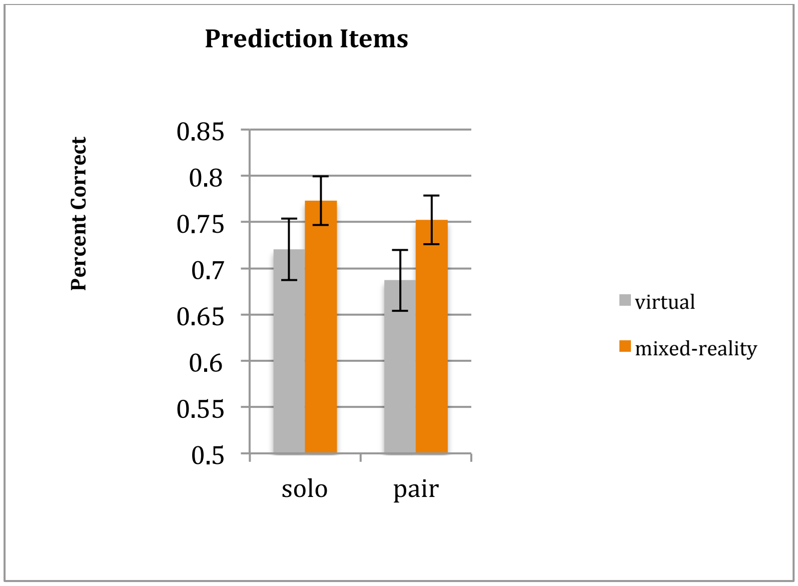



**Figure 6b.** Results for Post-test Prediction Items

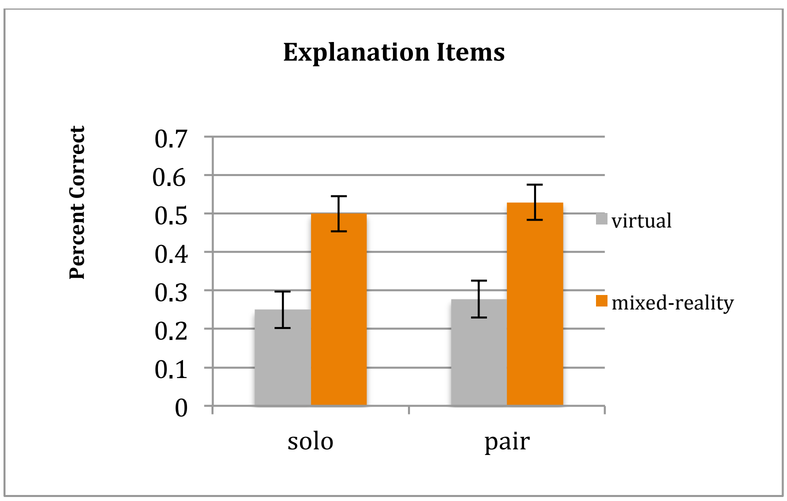



**Figure 6c.** Results for Post-test Explanation Items

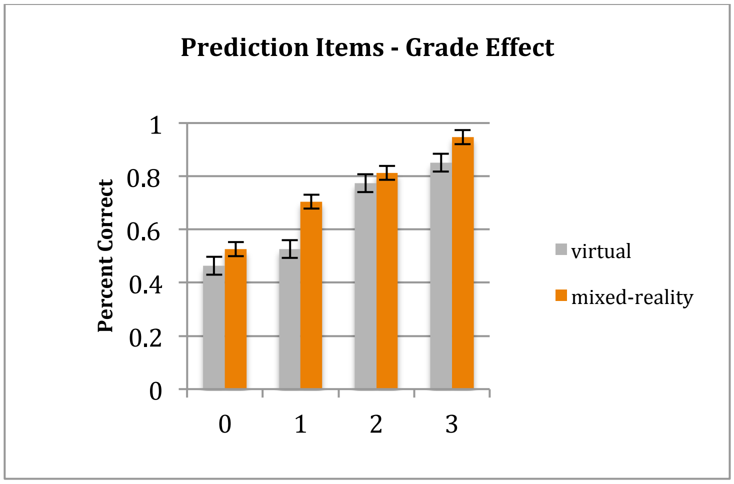



**Figure 6d.** Prediction Items Grade Effect

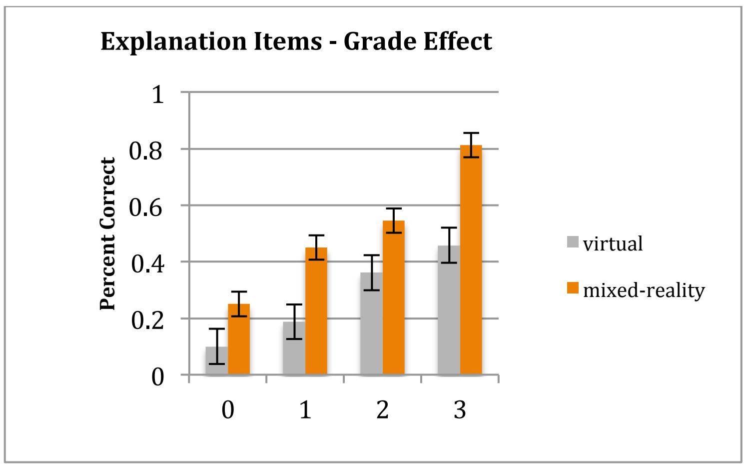



**Figure 6e.** Explanation Items Grade Effect

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

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

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

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

All three measures (the prediction items, the explanation items, and towers) showed a significant positive effect of mixed-reality conditions. What might explain this benefit? This

---

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

thesis explores three likely mechanisms suggested by prior work: collaboration, engagement, and embodied cognition [Carini et al. 2006; Antle 2013]. Comparisons of the solo and pair conditions did not suggest any effect of collaboration. My quantitative and qualitative analyses, described below, provide evidence for embodied cognition but not engagement.

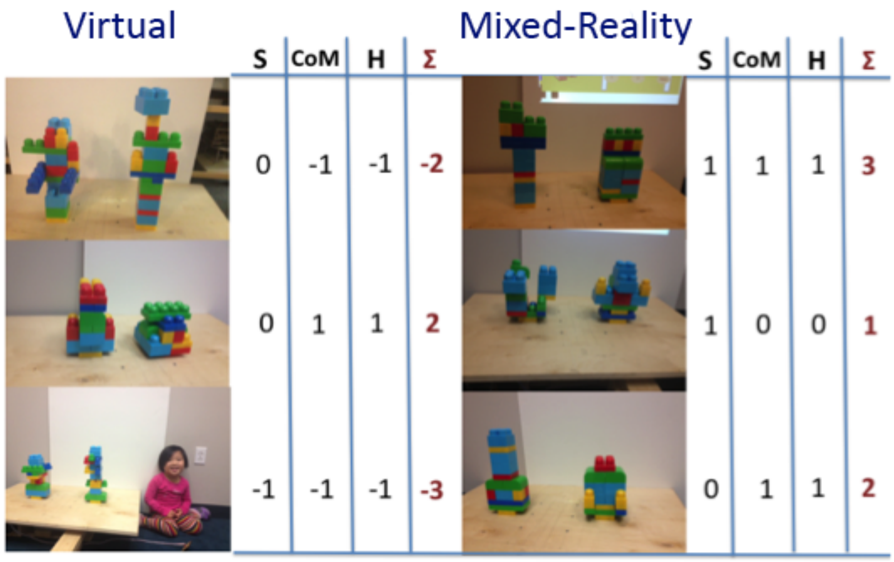



**Figure 7.** Coding scheme for Tower pre/post tests change.

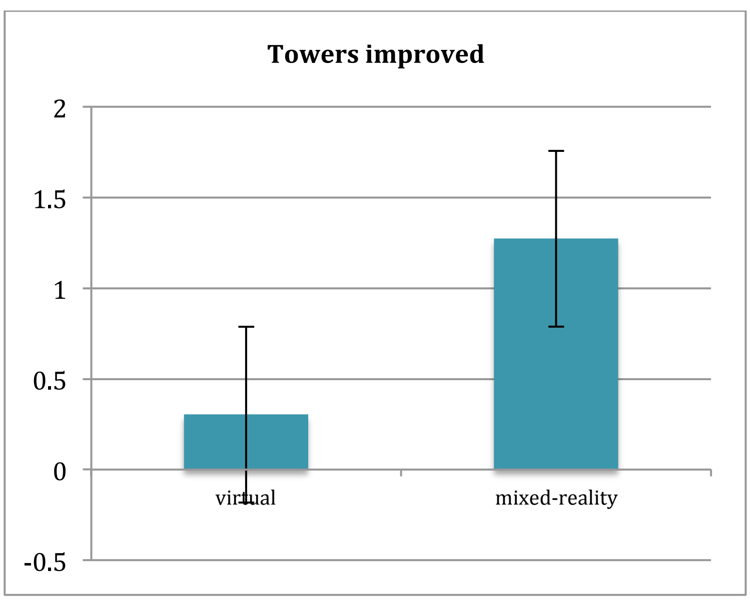



**Figure 8.** Scores on the tower building task (out of three). Positive scores indicate pre-to-post improvement in stability; a score of 0 indicates no change from pre to post.

### 4.3 Engagement and Enjoyment

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

**Figure 9.** Engagement and excitement of children when they see that their prediction was right.

The formal survey (given after the post-test) provides another measure of enjoyment. Students were asked how much they liked the game, and responded with options on a likert scale 1-5 ("I didn't like it at all", "I didn't like it", "It was OK", "I liked it", "I liked it very much"). Students in the mixed-reality condition had higher mean ratings for enjoyment, and an ANOVA showed that this difference was significant ($F(1,66)=6.9$, $p=0.01$, $d=0.48$). There was no significant difference between the solo and pair groups for likability (See

Figure 9. 1-5 scale was converted to proportion 0-1). However, the difference in enjoyment does not explain the difference in learning. Repeating the learning analysis only for students who gave the highest enjoyment rating still indicated a significant, favorable effect of mixed-reality (p=0.001).

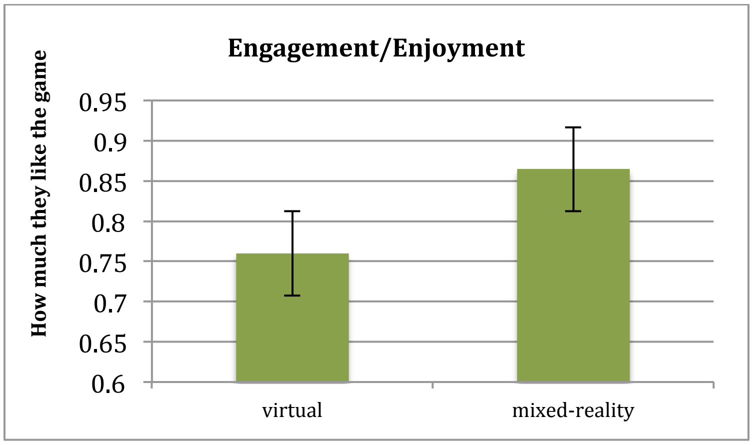



**Figure 10.** Results of the survey given to measure how much children enjoyed the game.

### 4.4 Gestures as Signs of Embodied Cognition

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

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


**Figure 11.** Children in the mixed-reality condition (above) used more shape-relevant gestures while explaining their predictions than those in virtual condition (below).

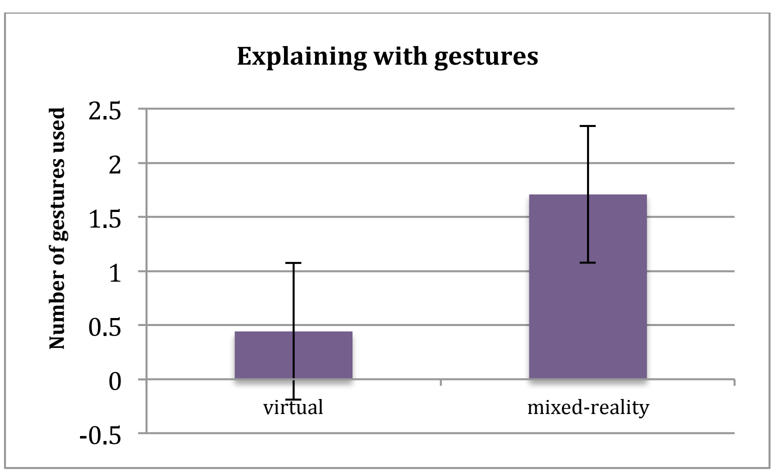



**Figure 12.** Average number of meaningful gestures used to explain predictions.

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

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

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

### 4.5 Qualitative Evidence

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

Most of the a-ha moments occurred after children made a wrong prediction and then recognized the relevance of one of the explanation options. For example, one child predicted that the left tower (Figure 4 – contrasting case 8) would fall first. Once the table shook, she saw that her prediction was wrong. When the multiple-choice explanation menu appeared on the display screen, she quickly selected her answer, exuberantly exclaiming

"Ooooh because it has more weight on top than bottom!" I suspect that observing the physical outcomes rather than the video leads children to take evidence against their prediction more seriously and thus more actively engage in trying to find an explanation. One child in each condition commented that they would prefer seeing the towers fall in real life rather than having a video or the computer say what happens.

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

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

# Chapter 5.
# Experiment 2: Mixed-Reality and Physical Control

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

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

## 5.1 Procedure

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

### 5.1.1) Mixed-reality version of EarthShake with Mouse Control

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

### 5.1.2) Mixed-reality version of EarthShake with Physical Control

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

**Figure 13.** Students using a physical switch to shake the table while interacting with the mixed-reality version of EarthShake with physical-control.

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

**5.1.3) Screen-only version of EarthShake with mouse control**

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

**Figure 14.** Virtual & physical-control condition where children shake the tablet to shake the table on the screen.

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

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

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

**5.2 Results of Experiment 2**

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

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

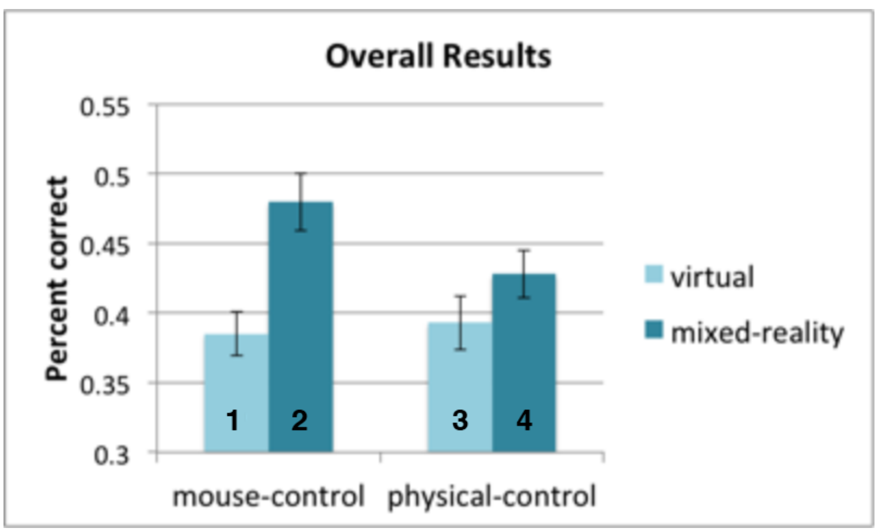



**Figure 15.** Overall post-test results.

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

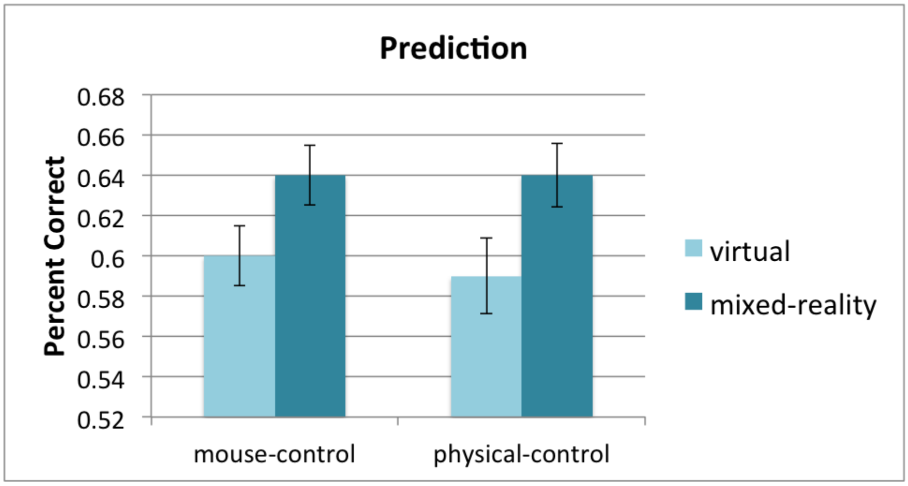



**Figure 16.** Post-test scores for Prediction Items

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

Likewise, for the explanation items, a 2-way ANCOVA showed significant differences in learning by media type, with the mixed-reality condition scoring higher at post-test than the virtual condition (Figure 16; 27% vs. 18% for 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. While the interaction between control type and media type is not significant, we do observe a trend: for students with the mixed-reality game, the, mouse-control condition was slightly better than the physical-control condition. One explanation for this trend could be that pressing the physical switch was so exciting for the children (supported by the data in the enjoyment section below) that they did not pay full attention to the explanations provided in the game.

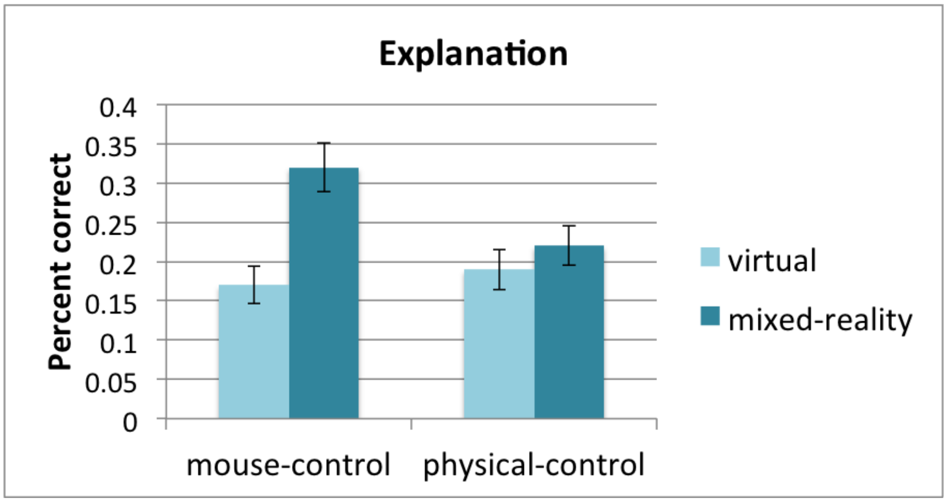



**Figure 17.** Post-test scores for Explanation Items

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

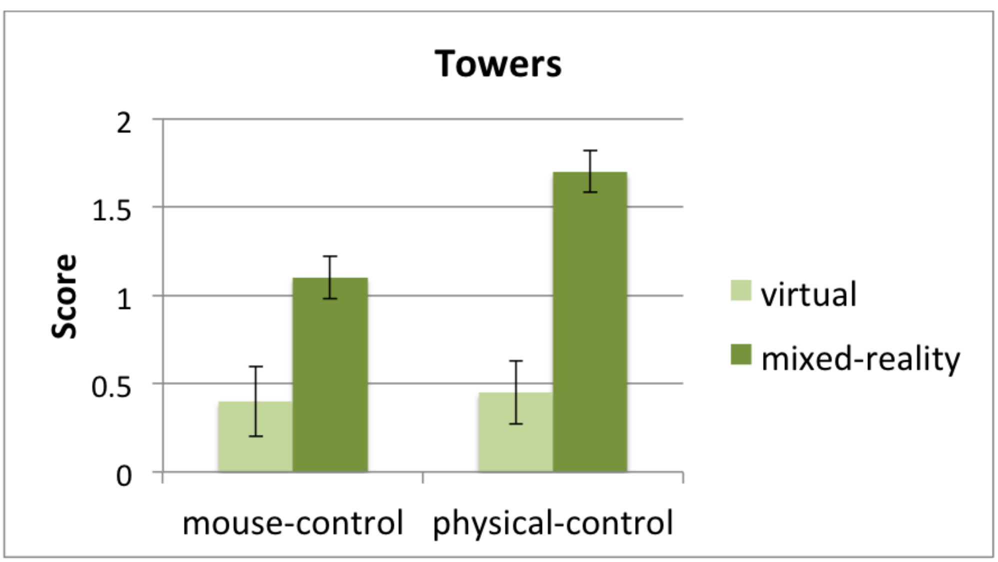



**Figure 18.** Tower scores

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

### 5.3 Engagement and Enjoyment

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

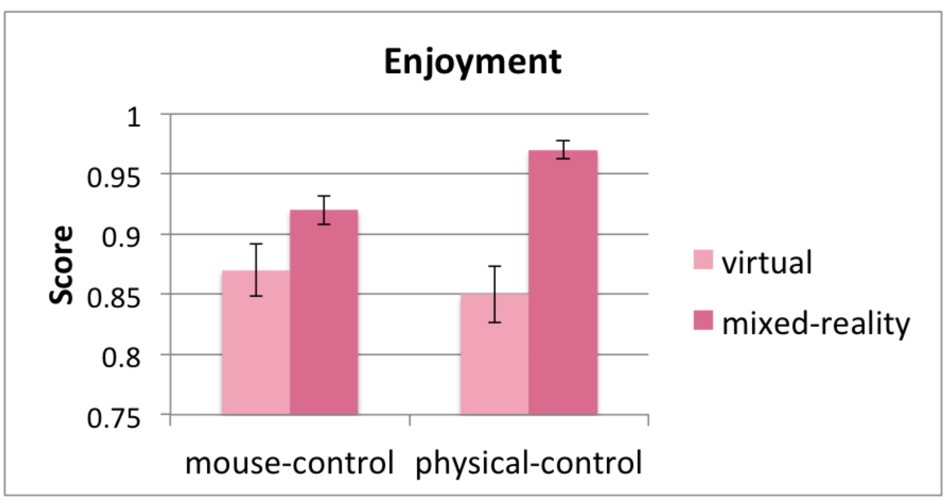



**Figure 19.** Enjoyment scores based on the survey

### 5.4 Gestures and Signs of Embodied Cognition

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

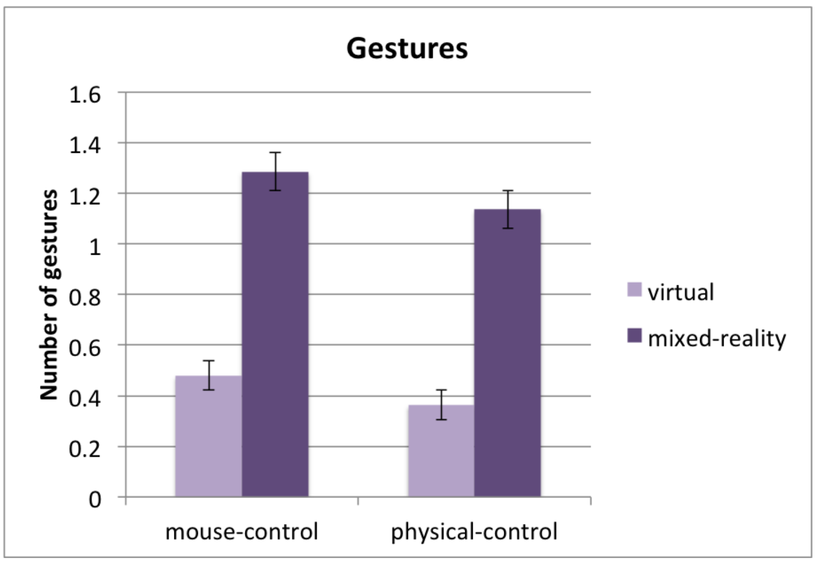



**Figure 20.** Meaningful gestures children used during their explanations.

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

### 5.5 Qualitative Evidence

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

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

# Chapter 6. Discussion of Experiment 1 and 2

EarthShake uses an affordable camera and a projector to combine the advantages of the physical and virtual worlds. In doing so, it presents a new kind of mixed-reality learning that incorporates prediction, physical observation, explanation, and personalized immediate feedback. Experiment 1 revealed significant differences in learning between the virtual and mixed-reality conditions, as measured by greater pre to 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, this finding begins to tease apart the effects of observing versus manipulating physical objects – factors that have not been controlled for in previous work. Experiment 2 replicated this intriguing result and also showed that a simple hands-on control, such as shaking a tablet or pressing a switch, does not have a significant effect on learning or enjoyment.

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

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

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

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

The gesture data provides some support for the explanation that physicality supports embodied cognition and triggers affordance for action, which as a result helps children perceive, mentally visualize, and ultimately remember concepts better. Children in the mixed-reality conditions more often explained their predictions using meaningful gestures to show 3D motion than children in the virtual condition. This finding suggests that those students had mentally visualized the objects, which may have helped them register and remember the explanations for why each tower fell. This result is in line with prior work that suggests (1) when children learn abstract concepts, they utilize mental simulations based on concrete motor-perceptual experiences [Antle 2013]; and (2) the gestures children spontaneously produce when explaining a task are a sign of their mental visualizations and predict how much they will learn from that task [Cook et al. 2008].

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

# Chapter 7.
# Design Implications from Experiment 1 and 2

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

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

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

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

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

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

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

# Chapter 8. Guided Discovery and Exploration in a Mixed-Reality System

Our previous research suggests that having a mixed-reality environment bridging physical and virtual environments produces enhanced learning and enjoyment. However it is not clear, from previous literature, what features of a mixed-reality system are most important for maximizing learning and enjoyment and, in fact, many researchers advocate for features that are at odds with the current design of EarthShake. As mentioned above, most tangible interfaces and mixed-reality environments to date have been purely exploratory and there aren't enough experiments that investigate how they should be designed to enhance learning. Also with the Maker Movement becoming more and more popular in today's world, there is a lot of emphasis on exploration with physical objects and exploratory engineering tasks, however the importance of guided discovery in such environments has not been explored much. For example, most Children and Science Museums have exhibits that encourage exploration with physical objects without much guidance. Our mixed-reality technology, provides the ability to add guided-discovery with personalized interactive feedback while children are doing physical exploration and experimentation, thus I explore the role of exploration and guided discovery in a mixed-reality setting.

There is research, which suggests that tangible interfaces are best suited for exploratory activity and learning through a process of discovery [Marshall et al. 2010]. One example of a tangible interface that affords exploratory activity is Underkoffler and Ishii's Illuminating Light [Underkoffler and Ishii 1998], which is designed to enable the rapid prototyping of optical layouts. Users of this optical prototyping tool move physical representations of various optical elements about a workspace, while the system tracks these components and projects back onto the workspace surface the simulated propagation of laser light through the evolving layout. Another tangible interface that encourages exploration is BitBall: a transparent, rubbery ball (about the size of a baseball) with a Cricket (a microcontroller-based electronics package), accelerometer, and colored LEDs embedded inside. Users can throw the ball up in the air and see the changing acceleration of the ball as changing colors as it goes up and down [Resnick et al. 1998]. Chromarium is a mixed reality activity space that uses tangibles to help children aged 5-7 years experiment and learn about color mixing [Rogers et al. 2002]. A number of different ways of mixing color were explored, using a variety of physical and digital tools. For example, one activity allowed children to combine colors using physical blocks, with different colors on each face. By placing two blocks together children could see the combined color and digital animations on an adjacent screen with immediate visual feedback. In another activity children could drag and drop different colored digital discs and see the resultant mixes. A third activity allowed children to use the digital interactive surface, with a digital image triggered a physical movement on an adjacent toy windmill. Rogers et al. (2002) found that the coupling of a familiar physical action with an unfamiliar digital effect is effective in causing children to talk about and reflect upon their experience [O'Malley and Fraser 2004].

It is often claimed that tangible interfaces are particularly good for exploratory learning, as interaction with tangible systems is found to be more natural or intuitive to students than other types of interface, affording a particularly suitable environment for rapidly experimenting. Many believe that if you just let kids explore with physical materials, they will learn. However, little comparative work has been carried out, and it remains unclear which elements of tangible interface designs are critical in supporting learning activities [Marshall et al. 2010].

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

### 8.1 Learning Theories behind Guided Discovery and Exploration

Pure discovery and exploration learning is based on (some interpretations of) constructivist learning theory, where open-ended discovery learning and hands-on exploration is believed to lead to better learning [Papert 1980]. Based on inquiry, discovery learning expects the learner to construct his/her own learning agenda. By experimenting and wrestling with results, the learner constructs new learning based on experience. Some have argued that pure discovery leads to better learning compared to guided instruction in Tangible User Interfaces [Schneider et al. 2015]. Schneider et al. built EarExplorer, an interactive tangible system where students can manipulate and connect parts of the auditory system to rebuild a functional structure. An augmented reality layer displayed sound waves

and showed how they are transformed at different stages of the process. Their previous work suggests that TUIs are particularly good at preparing students for future learning; that is, students learn more when they can explore a novel domain with a TUI before compared to after receiving a traditional (e.g. lecture or
text based) instruction. With a study on EarExplorer, they aimed to isolate the impact of structured guidance versus no guidance during a hands-on TUI activity on learning. In one condition in the study, students rebuilt the hearing system by self-driven discovery, while in another condition they rebuilt it by following the step-by-step instructions of a video-teacher. They found that the first group ("discover") significantly outperformed the second group ("listen") by ~27% on the final learning test [Schneider et al. 2015].

Kirschner, Sweller and Clark criticize that discovery learning defies the cognitive load theory – the learners cannot retain the amount of information needed to process the content [Kirschner et al. 2006]. Other critics have also pointed out the possibility of undetected misconceptions, student frustration and lack of worked examples [Tuovinen and Sweller 1999]. While Papert (1980) argued that by using Logo children will have "mindstorms" and acquire "powerful ideas" - that was the dream, but not the reality. Students do not learn powerful ideas from Logo [Pea and Kurland 1984], unless the activity context is well engineered and targeted at well-defined learning objectives [Clements 1990; Clements 1986; Klahr and Carver 1988]. The issue of activity context and the relation to educational objectives is even greater in educational games where there is the potential for students to be distracted by the game goals and, thus, not achieve the learning goals [Miller et al. 1999].

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

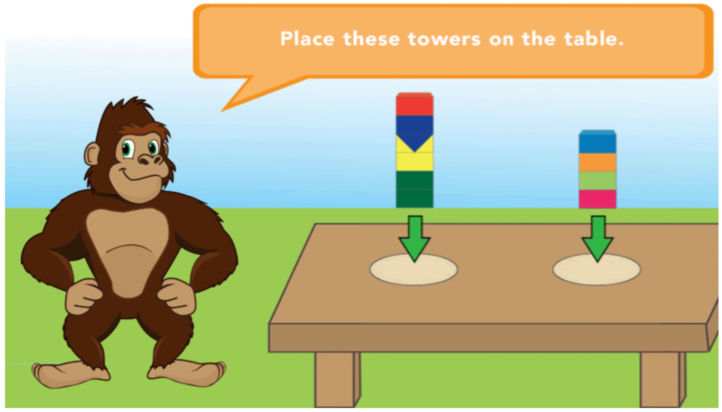
Miller et al. suggest that the learning outcomes achieved through microworld interaction depend largely on the surrounding instructional activities that structure the way students use and interact with microworlds [Miller et al. 1999]. In an experimental study they compared the standard-goal and specific-path approaches with a third no-goal condition and found that students in the standard-goal condition generally learned less qualitative physics than those in the two alternative conditions. They propose that the no-goal condition and specific-path predisposed students to scientific modeling, whereas the standard-goal condition predisposed students to an engineering approach. They saw superior performance by the no-goal group on the scientifically-oriented post-test. Their results suggest that careful selection and analysis of the tasks that frame microworld use is essential if such environments are to lead to the learning outcomes imagined for them [Miller et al. 1999]. To take a broader view of the issue, these results are consistent with Schauble et al.'s findings where they distinguish between students who use an engineering model of experimentation and those who use a science model [Schauble et al. 1991]. The behavior of the engineering group was characterized by manipulation of variables to produce a desired outcome, whereas the science group was characterized by broader exploration and more selectiveness in interpreting evidence, especially disconfirming evidence. They investigated the hypothesis that when children are engaged in science experiments, the goal of which is to understand relations among causes and effects, they often use the engineering model of experimentation, characterized by the more familiar goal of manipulating variables to produce a desired outcome [Schauble et al. 1991].

A related set of invited hypotheses that the proposed experiment explores here are: 1) whether or not an exclusive focus on activities oriented toward engineering thinking might limit outcomes, especially scientific learning outcomes, 2) whether or not an exclusive focus on activities oriented toward scientific thinking might limit outcomes, especially engineering learning outcomes, and 3) whether or not a mix of activities produces both types of outcomes and perhaps as well even with less time devoted to each activity type. Based on these hypotheses, I aim to discover if guided-discovery with scientific thinking practice, exploration with engineering practice or a combination of both leads to better learning outcomes in mixed-reality environments.

### 8.2 New Design of EarthShake

In order to test the effect of exploration and guided discovery in a mixed-reality environment, I have decided to add more exploratory learning activities to EarthShake. My qualitative analysis from the experiments suggest that it may have benefits to have more hands-on activities in the game where children get to manipulate the physical objects and do more hands-on exploration (supported by Montessori's theory that young children are highly attracted to sensory development apparatus and that they use physical materials spontaneously, independently, and repeatedly with deep concentration [Montessori 1964]. My gesture analysis, showed that children perform mental visualizations when they experiment/observe physical objects, leading to better learning. So it may be argued that having more hands-on manipulation and building activities with physical objects may lead to even more mental visualizations and thus better learning. From the surveys, I also saw that kids enjoyed building their own towers and testing them on the earthquake table. Thus based on the feedback and literature about hands-on learning, I decided to add more hands-on activities into the game. The new design of the game consists of two different modes: 1) Guided Discovery with Predict/Observe/Explain/Feedback cycle, 2) Hands-on Exploration and Problem Solving. It starts with a main menu, where users can choose either to play the game, or build their own tower. If they choose "Play the Game" option then they go directly to the guided discovery activity where they are asked to place two towers on the table and make a prediction about which one will fall first. On the other hand, if they choose "Make a Tower" option, then they go to the hands-on exploration activity where they can build their own tower and test how long it will stay when the earthquake table shakes. They can go back and forth between different modes, and are asked if they want to go to the other mode after interacting with the game a few times.  In other words, "Guided Discovery" mode is designed to encourage more scientific thinking practice, while "Exploration" mode (where they're asked to build their own towers that won't fall down when the table shakes) is designed to be more aligned with the engineering approach introduced in the Learning Theories section above.

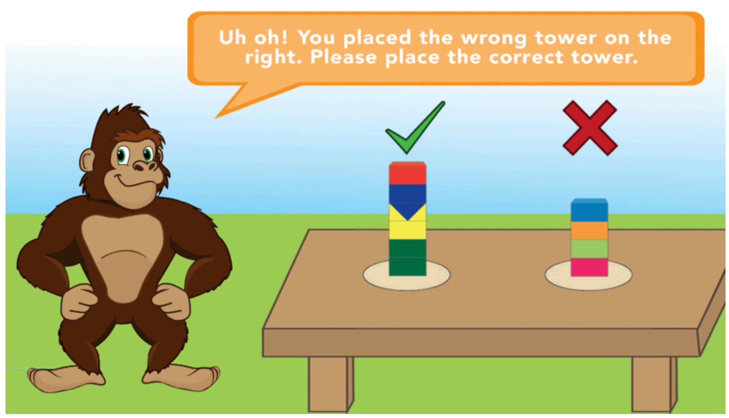
**a)**


Place these towers on the table.

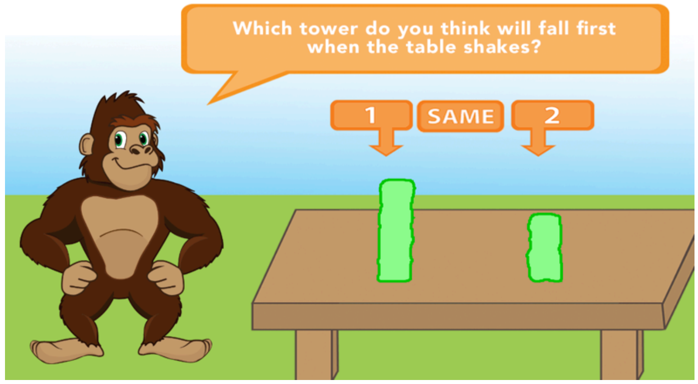
**b)**


Good job! Click to continue.
CONTINUE

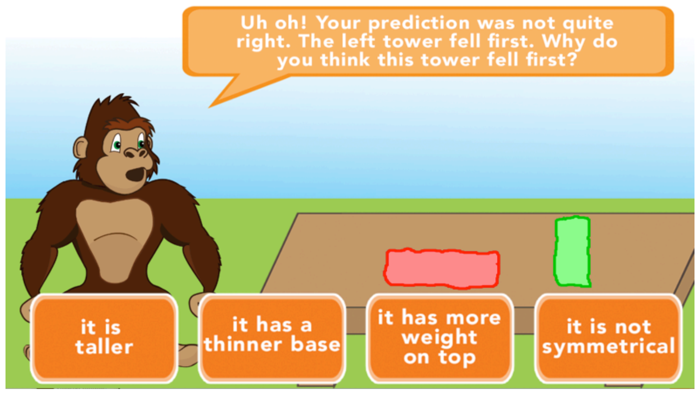
**c)**


Uh oh! You placed the wrong tower on the right. Please place the correct tower.

d)




e)




f)


**Figure 21.** Screenshots from the Guided-discovery mode of the game. Users are guided to place the given towers on the table (different than the older version of the game), and are given feedback whether they placed the right tower or not. The game also gives feedback about their explanations with visualizations to help them understand the underlying physics principles.

### 8.2.1 Guided Discovery with Predict/Observe/Explain/Feedback

The "Guided-discovery mode", takes advantage of the predict-observe-explain cycle with personalized interactive feedback that is provided through a gorilla character in the game. This mode is illustrated in Figure 21. First, the users are asked to place the towers shown on the screen on the physical earthquake table (See Figure 21a). The Kinect camera and our specialized computer vision detects if the tower placed on the table matches the one on the screen (based on the moment of inertia based values in the database determining the shape of the object as explained in Section 2.2). If the tower they place on the table matches the tower on the screen, a check appears on the tower on the screen and the gorilla character says "Good job! Click to continue" (See Figure 21b). Otherwise, if the tower they place does not match the tower on the screen that they were asked to place, the computer vision system detects that it wasn't the correct tower and a cross appears on the tower on the screen and they're asked to place the correct tower (See Figure 21c).

Once they click continue, the gorilla character prompts them to make a prediction about which of the two towers on the table will fall first when the table shakes, by saying "Which tower do you think will fall first when I shake the table?" (See Figure 21d). They can choose either 1, 2 or same by clicking on the buttons or one of the towers on the screen (a virtual projection of the towers as a blob is drawn on the screen).

After they make a prediction by choosing which of the towers they think will fall first, the gorilla character says: "You chose the left tower. Why do you think so? Discuss and then click SHAKE to see what happens". Here they can discuss their prediction with their partner/friends/family, why they think the tower they chose will fall first. They can click the "Shake" button to shake the table and observe the results. When they click the "Shake" button on the tablet, the physical earthquake table starts shaking (It triggers the relay that is connected to the motor of the earthquake table).

After the table starts shaking, when one of the towers falls down, the Kinect camera and our specialized computer vision algorithm detects the fall (based on the height of the detected blobs) and stops the earthquake table from shaking. The vision algorithm detects whether the left or right tower fell and if it matches with the prediction of the user. If the user's prediction was correct and the right tower fell first, then the gorilla character says "Good job! Your hypothesis was correct! Why do you think this tower fell first?". The gorilla on the projector screen is happy and starts jumping and dancing to give them positive feedback and reward. On the other hand, if the user's prediction was wrong and the tower that fell does not match the tower that was predicted by the user, then the gorilla character says "Uh oh! Your prediction was not quite right! Why do you think this tower fell first?". This time, the gorilla on the projector screen is sad and surprised. Then they are asked to explain why they think the tower that fell, fell first. This time, they are given an explanation menu with four different choices that they can choose from: "It is taller", "It has a thinner base", "It has more weight", "It is not symmetrical". These are the four different principles of stability and balance. They can click on one of the answers to explain the reason for the results they have observed.

When they click on one of the choices in the explanation menu, the gorilla character tells them if their explanation was correct or wrong, with a visualization laid over the images of the towers on the screen to explain why the tower actually fell (See Figure 21f). For example, if the reason was because it had more weight on top than their bottom, and their explanation was not correct, he says: "Actually it fell first because it had more weight on top than bottom. Good try. Click CONTINUE to play again." And the visualization of the towers shows circles on the parts of the towers that have more weight. Similarly, if the tower fell because it had a smaller base, the visualization shows ruler visualizations on the towers, showing that the width of the bases is different in each tower. If it fell because it was taller, the ruler visualizations this time highlight the height of each tower. Finally, if the tower fell because it was not symmetrical, there is an overlay on the towers, which shows a dotted line that splits the tower into two pieces showing if it's the same on each side or not.

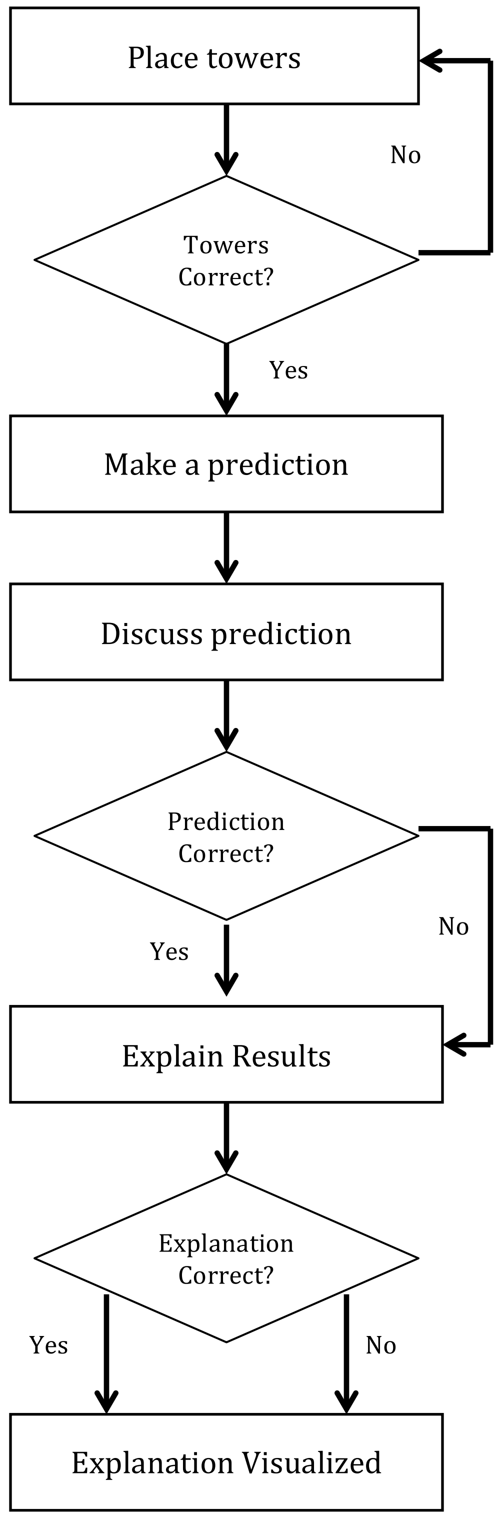
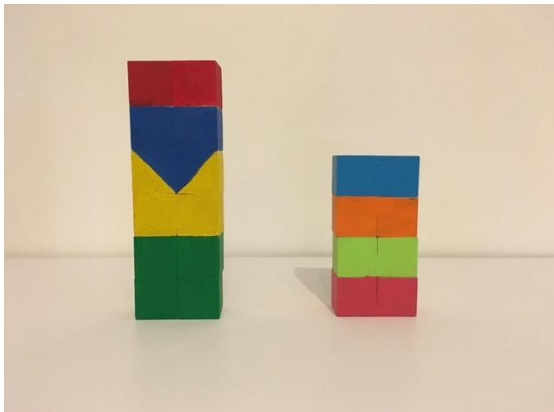
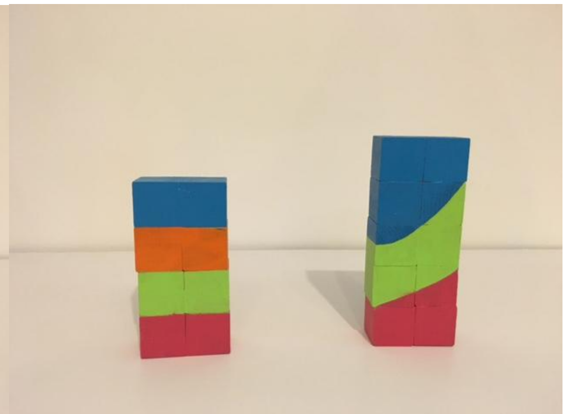



**Figure 22.** Activity Diagram

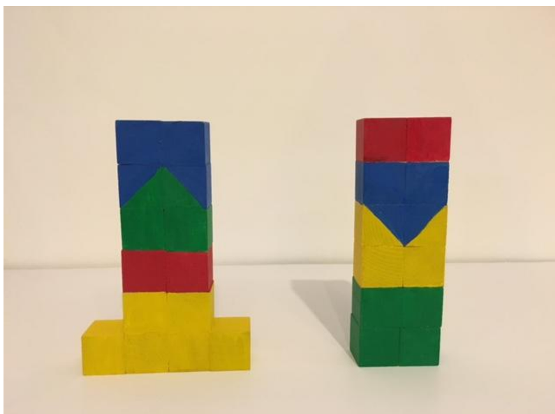
This scenario is repeated for different contrasting cases.



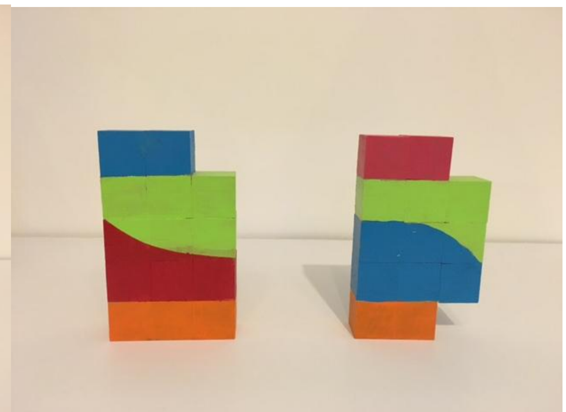
HEIGHT (D2, D3)



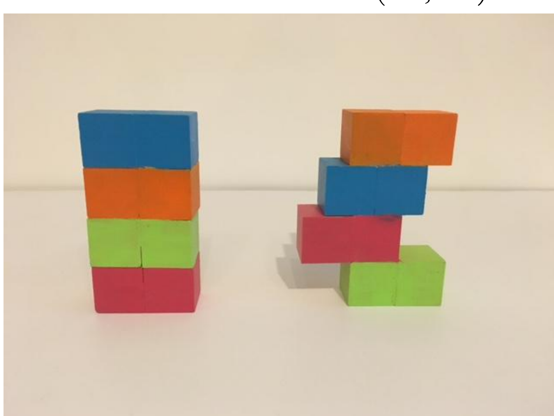
HEIGHT (D3, B1)



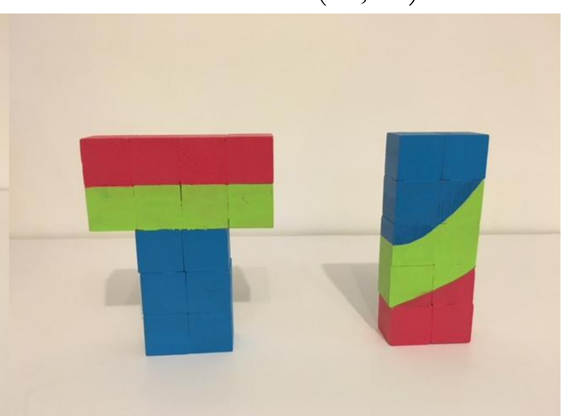
WIDE BASE (D1, D2)



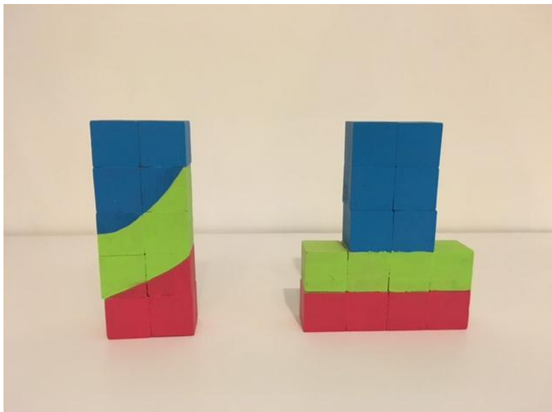
WIDE BASE (B4, B3)

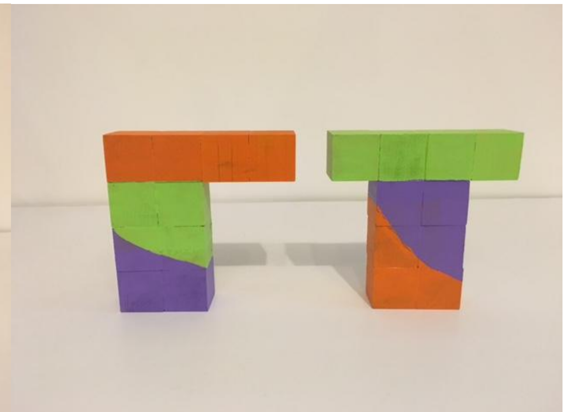


SYMMETRY (D3, D4)

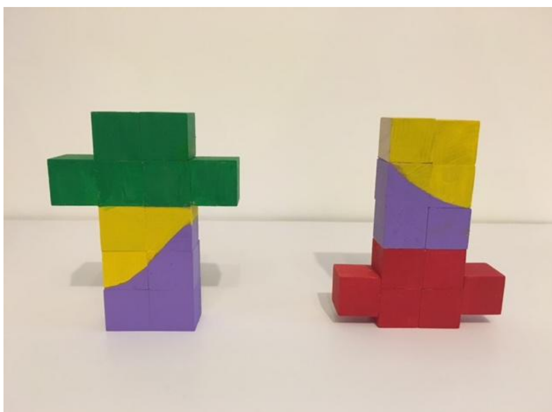


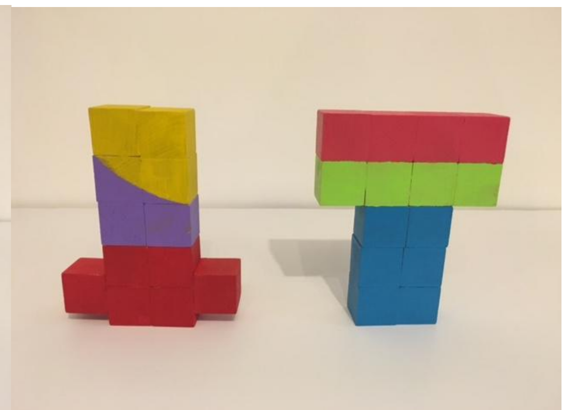
CENTER OF MASS (E2, B1)





WIDE BASE (B1, B2) SYMMETRY (C1, C2)





CENTER OF MASS (A1, A2) CENTER OF MASS (A2, E2)

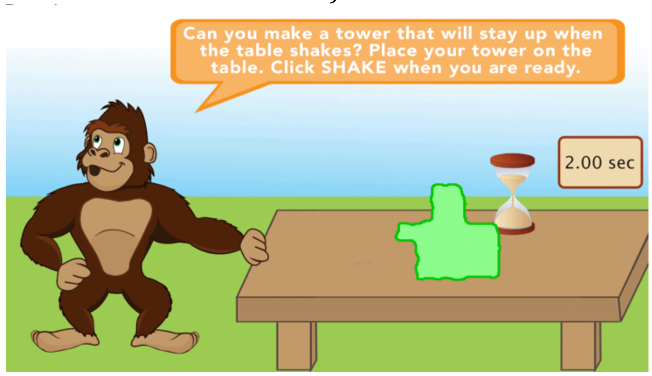
**Figure 23.** Contrasting case towers given in the game to teach different principles.

### 8.2.2 Hands-on Exploration and Problem Solving

In the "Hands-on Exploration and Problem Solving mode", the users can build any tower they want using the wooden, Lego or magnetic blocks in the bins. The gorilla character says: "Can you make a tower that will stay up when the table shakes? Place your tower on the table and click SHAKE when you are ready." (See Figure 24 a) The only restriction is that they have to build one tower to be tested on the earthquake table. If they place more than one tower, they are prompted by the gorilla character to make sure there is only one tower on the table: "Oops! There are too many towers on the table. Please make sure there is only one tower and then click SHAKE to see what happens." When they have built their tower and click the SHAKE button on the tablet, it triggers the motor in the physical earthquake table, and the table starts shaking. If their tower falls down, the Kinect camera and our specialized computer vision algorithm detects the fall, and the gorilla character gives feedback "Uh oh! Your tower fell down! Press CONTINUE to make another tower" (Figure 24 d). The system also displays how many seconds it took for the tower to fall down. If the tower does not fall down in 5 seconds, the earthquake table stops shaking. Then the gorilla character starts jumping/dancing (similar to that in the guided-discovery mode), and says: "Good job! Your tower stayed up! Press CONTINUE to make another tower" (Figure 24 c). They can then make another tower to test on the earthquake table.

a)


b)



c)


d)


**Figure 24.** Screenshots of the Exploration mode

### 8.3 Piloting at the Children's Museum of Pittsburgh

I created the new design of EarthShake including the guided-discovery and exploration modes and then had two play testing sessions at the Children's Museum of Pittsburgh to get feedback from families and children and iterate on the design based on the feedback.

#### 8.3.1 Pilot 1

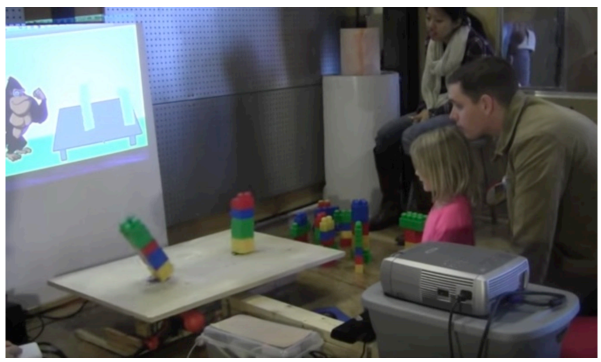
##### 8.3.1.1 Method

The first pilot was took place in a separated room of the MakeShop of Children's Museum of Pittsburgh, which allowed us to have more control and get more feedback from the participants. There was a sign at the door of the room where the pilot was done, which said that only one group/family at a time can come in. The pilot took place over two days in the weekend, from 10 am until 5 pm when the museum closed. Around 40-50 people interacted with the exhibit over 2 days.

#### 8.3.1.2 Qualitative Data and Findings from Museum Pilot 1

It seemed like the parents were very much involved in helping the kids. At some points, they even started playing the game themselves (See Figure 22). The kids also seemed to be very much engaged. A child who had a birthday party that day, played with the game early in the morning before anyone else came, then she brought her whole birthday party to play later on. Some of the parents were also curious about how the system worked. For example, one of the dads said: "Very cool!" and asked how we had built the system and if he could set it up at home.

*Usability/design issues:* One issue I found was that the younger kids had a hard time using the mouse. Therefore I decided to use a tablet as an input device instead of the mouse. Also, our computer vision algorithm that was dependent on the colors of the blocks seemed to be effected by the lighting conditions in the room. Therefore, I decided to change the algorithm so it would use depth and shape information instead of color. We came up with a new algorithm that calculates the moment of inertia based values of towers using their shape to identify which tower has been put up (explained in Section 2.2). Also, I saw that the prebuilt towers that I had created by sticking Lego blocks together were not durable. Children thought that they could be taken apart when they saw the Lego blocks (glued together to create prebuilt towers) and tried to separate them. However, I realized that the colors of the towers were actually very helpful for kids and parents to distinguish between the towers and use as cues to discuss/decide which tower to put up. Thus for the future experiments, I decided to create prebuilt towers made of wood blocks that would be secured tightly to each other and painted to create a visual cue. However the Lego blocks were still used for the exploration mode where children could build any tower they wanted using Lego blocks, since they did not need to be glued/prebuilt for this case.



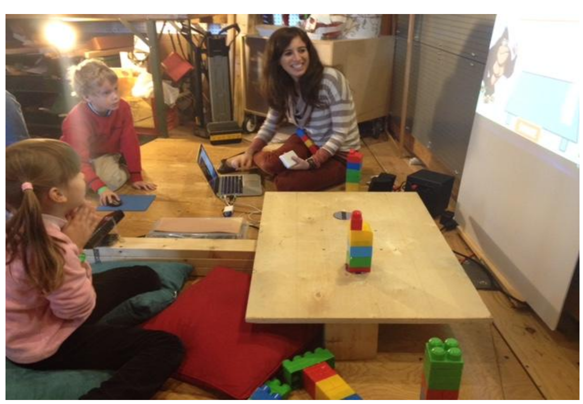


**Figure 25.** Photos from the first play testing at the Children's Museum of Pittsburgh. Both children and parents seemed to be quite engaged in the exhibit.

**8.3.2 Pilot 2**

I iterated on the design based on the feedback from the first pilot, and came up with a new design. For this pilot and subsequent experiments, a second iteration of the physical setup was used. This setup had a more robust mechanism and more polished physical design (See Figure 27).

**8.3.2.1 Method**

The second pilot was took place in the MakeShop of Children's Museum of Pittsburgh in a less-constrained area open to public, which allowed us to test our system in a real world setting. There was a sign explaining our research. The pilot took place over two days in the weekend, from 10 am until 5 pm when the museum closed. Around 40-50 people interacted with the exhibit over 2 days.

When they approached the exhibit, there was a screen where they could choose one of the two options: "Play Game" or "Test my Tower". If they chose "Play Game" option they started interacting with the Guided-discovery mode of the game. On the other hand if they chose "Test my Tower", then they started interacting with the Exploration mode. After 4 contrasting cases, they were prompted to a transition screen, where they were asked if they would like to continue playing or build their own tower. Below are some screen shots from the pilot to demonstrate the sequences they went through:


**Figure 26.** Users were asked to place the towers on the table (shown on the game screen).

**Figure 27.** Users were asked to make a prediction about which of the towers will fall first.


**Figure 28.** Users were asked to discuss why they think the tower they chose will fall down. They can then click the "Shake" button on the screen to shake the physical earthquake table.

**Figure 29.** Users observed the results with interactive feedback from the system.


**Figure 30.** The system gave feedback about their explanation, if their explanation was right or wrong, with visualization laid over the images of towers to explain why the tower actually fell.


**Figure 31.** Children could build any tower they want using the wooden blocks or lego blocks. The system gives feedback about how many seconds it took for their tower to stay up. If the tower that the user built stays up, the gorilla character made his little dance. Otherwise, if their tower fell down, the gorilla character was surprised and an hourglass was displayed showing how many seconds it took for the tower to fall down.


**Figure 32.** Children imitated the gorilla character when their prediction was correct and the gorilla started dancing on the screen. The parents were also very much engaged with the game. At some points they started playing on their own.

#### 8.3.2.2 Qualitative Data and Findings from Museum Pilot 2

I had some interesting observations from the pilots. Both kids and parents seemed to be very much engaged by the exhibit. I observed that parents and children seemed to discuss and engage in productive dialogue, discussing the reasons why the towers fall, more in the guided-discovery condition. When the game prompted them to explain why they think one of the towers fell first, the parents started asking children why they think the tower fell and helped them understand the underlying principles. For example, one of the moms put the towers side by side, asking her son what the difference is between the two, if one has more weight on top than the other. Similarly, many other parents got involved and tried to guide the kids to understand why. One of the moms made high five with her daughter when their prediction was correct and their tower stayed up (Figure 20).

In some cases the parents even started playing on their own. One man played the game for a while (around 15-20 minutes), trying to build a tower that stays up. At the end when his tower finally stayed up, he raised his hands showing his pride that his tower had stayed up (Figure 20). Also, children seemed to be engaged by the gorilla character. For example, when they made a prediction and their prediction was correct, the gorilla character started dancing on the screen. Some of the kids started imitating the gorilla character as she danced on the screen, dancing along with it.

I also observed that children and families seemed to strategize more while building their towers when they did it after the guided discovery activity as opposed to doing it as a first activity when they first came to the exhibit. If they chose to build a tower before interacting with the guided-discovery activity, they tended to do more random things that are not aligned with the goals of the game, e.g. trying to make a tall tower that will fall quickly. These observations informed the research questions and hypotheses for the experiment below (i.e. do people learn better from exploration or guided discovery or a combination of both?).

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

I also discovered some design issues that I iterated on before the experiments. For example, the tablet was inserted in a hole on the wooden table. However the kids tried to pull the tablet from its place, which caused the connection to the computer to be lost. So I decided to make a wooden frame on top of the tablet that would hold it in place so that they couldn't pull it away.

## 8.4 Experiment 3: Guided Discovery vs. Exploration/Problem Solving

### 8.4.1 Experimental Design

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

These questions lead naturally to three experimental conditions as shown in Table 2. For the Guided Discovery condition (G), children were given 10 contrasting cases, for which they were asked to make a prediction, observe and explain the results with interactive feedback as explained in section 8.1.1. For the Exploration condition (E), children were asked to build towers that will stay up as long as possible when the table shakes. As explained in section 8.1.2, this condition involves limited interactive feedback that informs children if their tower stayed up or not and for how long it stayed up. They were asked to build multiple towers for a time period that is equivalent to the time required in the Guided Discovery condition. For the Combined condition (C), children were given 5 contrasting cases (targeting different principles), followed by an exploration task (building their own tower), and then 2 more contrasting cases followed by two more exploration tasks. Pilot testing was to inform the best combination and to be sure that learning time across conditions is well matched.

*Conditions:*

| Condition | Explanation |
|---|---|
| Condition 1: G | Guided Discovery (8.1.1) |
| Condition 2: E | Exploration/Problem Solving (8.1.2) |
| Condition 3: C | Combined Guided Discovery + Exploration (8.1.1 & 8.1.2) |

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

*Research Questions:*

The experiment helps us answer the main research questions:
How to make a mixed-reality environment that is optimized for learning and enjoyment:
1) Does guided discovery using predict-observe-explain-feedback structure help kids learn better compared to exploration/problem solving in a mixed-reality environment or is a combination of exploration/problem-solving and guided discovery better for increasing fundamental concept learning and enjoyment? (Condition 1(G) vs. Condition 2(E) vs. Condition 3(C)) If guided discovery activities are good for promoting scientific thinking, it may improve scientific outcomes such explanations in post-tests. If exploration and engineering approaches are better for fundamental principle learning, children who get to explore more (E condition) may have better outcomes on post-tests. On the other hand if combining engineering activities with scientific thinking practice leads to better engineering outcomes, we can expect better prediction results for the combined group (C).
2) Does guided discovery, exploration or a combination of the two transfers better to a hands-on tower building task? (Condition 1(G) vs. Condition 2(E) vs. Condition 3(C)). If exploration and practicing engineering activities are critical on their own for transferring to an engineering task (such as building a tower), we would expect those in the E condition to perform better. On the other hand if combining scientific thinking practice with engineering activities improves engineering outcomes, we would see that C condition would perform best on a tower construction post-test.

### 8.4.2 Participants

I conducted the study with 4-8 year old children in a lab or school setting with 75 children in first and second grades. Children were recruited from Montour School District. Children were randomly assigned to each condition and interacted with the game in pairs.

### 8.4.3 Procedure

The students were pulled out from their classroom in pairs. They came to a room in the library where the equipment was set up. Before interacting with the game, students were first given a tower pre-test, consisting of two tower tasks. First, the experimenter showed them a tower that was prebuilt (Figure 33) and told them: "I built this tower, but it is not very stable and it would fall down if I shake the table. Can you make a tower that is more stable using the same blocks" and handed them a bag of blocks that they can use to build a tower together with their partner. After they were done building their tower, the experimenter asked them to explain how they built it and if they had any strategies in mind. After they explained their towers, they were then given another bag of towers. This time they were asked to build a tower using all the blocks in the bag, using a specific red block on the bottom. The experimenter told them that there were two rules this time: first, they had to use all the blocks in the bag and they had to use the red block on the bottom and nothing else can touch the table. Again, after they built the tower in collaboration with their partner, they were asked to explain how they built their tower.

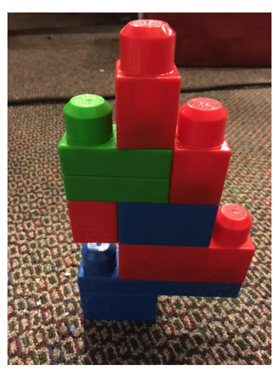


**Figure 33.** Participants were given a prebuilt tower and asked to make a tower that is more stable than this tower using the given blocks.

Then they were asked to complete a paper pretest to measure what they already knew about the stability and balance principles in the game. Students then interacted with their assigned game, either guided-discovery (G), exploration (E), or the combined guided-discovery and exploration (C) condition. The condition content was designed through piloting to involve the same time and some maximum time count-off was employed in the case some children take much longer than expected.

For the G condition, participants interacted with 10 contrasting cases (in the sequence given in Figure 23). For the C condition, participants interacted with 5 contrasting cases (in the following sequence: D2&D3, D1&D2, B4&B3, D3&D4, A2&E2), then they did one exploration activity in the game and then they interacted with 2 more contrasting cases and 2 more exploration activities (C1&C2 and A1&A2). This sequence was determined based on piloting before the experiment to match the timing of the G condition and to make sure children get exposed to enough contrasting cases before the first exploration. Thus the overall time on task was the same as G condition in the C condition, however the time spend on guided-discovery activities was less than the G condition (7 guided-discovery activities in C versus 10 guided-discovery activities in G). On the other hand for the E condition, they were only given exploration tasks where they were asked to build a tower that would stay up when the table shakes. If they were able to build a tower that stayed up, then they were asked to build a tower that is taller than their previous tower that would still stay up. Again, the time that children interacted with the game was calculated to match the other conditions.

After interacting with their game, students were given a matched paper post-test. After the paper post-test, the students were given the same tower building task as before game play. As in prior studies the contrast between the pre- and post- tower building assessment was used to measure student improvement in tower building, and in particular, how well they incorporate the principles of balance. Finally, the students were asked to fill out a survey to see how much they enjoyed the game.

### 8.4.4 Measures

The paper pre and post-tests are prepared based on the NRC Framework & Asset Science Curriculum [Quinn et al. 2012]. I used the same tests as I used in the previous experiments. The tests consist of two types of items: prediction items and explanation items. For the prediction items, the students were given a picture of a table with two towers and were be asked to predict which will fall when the table shakes. In the explanation items, they were be asked to explain why they chose their answer (Figure 21). Also, children were given a survey at the end of the game to measure their enjoyment.

The survey consisted of three questions (See Figure 34). The first question was: "How much did you like the game?" They could choose one of: "I didn't like it at all", "I didn't like it", "It was OK", "I liked it", "I liked it very much". The second question was: "Would you like to play it again?". They could choose "Yes", "No" or "Maybe" by choosing one of the smiley faces from a scale of 1-5. Finally, the third question was: "Would you recommend it to a friend?". Again, they could choose "Yes", "No" or "Maybe" by choosing one of the smiley faces from a scale of 1-5.

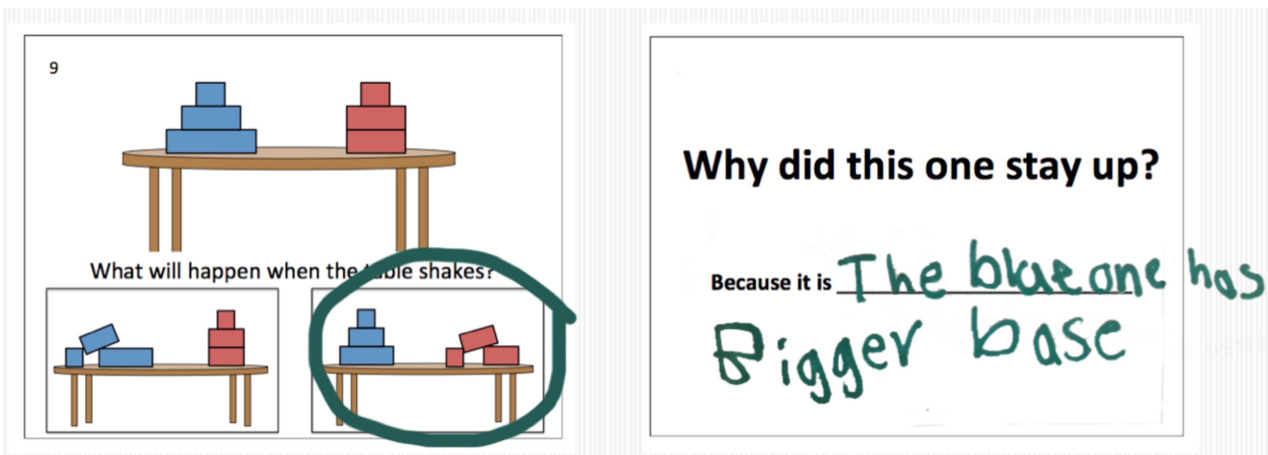



**Figure 34.** Prediction (left) and explanation (right) items used in the paper pre/post-tests.


**Figure 35.** Survey questions.

We also gave the participants some hands-on building tasks as pre and post tower tests, to see how much they could transfer the knowledge they gained to practical hands-on activities in the real world. Different than the previous experiments, this time we gave children two different tower tests, to have more hands-on activities as pre and post tests and have a better understanding of how their learning from the different modes of the game translate to real-world hands-on activities.

### 8.4.5 Hypotheses and Data Analyses

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

Perhaps one of the most straightforward claims about learning is that you learn what you practice. An associated strong instructional design recommendation is that instruction should be aligned with the learning goals as made concrete and observable through targeted assessment tasks (Carver et al., 2010). If you want students to learn to do tasks of a certain type (say type A), then give students practice on tasks of that type (type A) and do not give them practice on tasks of a different type (say, type B). Clark and Mayer (2011) provide a similar recommendation suggesting that practice (with feedback) should be designed to build job-relevant skills [Clark and Mayer 2016]. A straightforward application of this theory, leads to the hypothesis that children in the Guided Discovery condition (G) should perform better on the paper pre and post-tests than those in other conditions, because G is better aligned with the goals of the tasks on the paper pre/post tests. The tasks on the paper test assess children's ability to predict, observe and explain (those are the "job-relevant" skills in Clark & Mayer's terms) and the G condition involves practice on predicting, observing, and explaining. On the other hand, children in the Exploration (E) condition, by this theory, should perform better than child in other conditions on the Tower Building pre/post tests, since E provides the most practice building their own towers. The theory here is simple: if you practice tower building more, you learn tower building better. If you practice predicting and explanation more, you get better at predicting and explaining what will happen. This principle does not make strong predictions about differences in enjoyment in the different conditions.

A similar theory puts less focus on similarity in task types and more emphasis on similarity in underlying knowledge (facts, concepts, skills, principles) acquired in learning and transferred to assessments [Koedinger et al. 2012]. It also suggests that guidance is critical to effective acquisition. A consequential prediction is that guided discovery will produce better learning of the fundamental physics principles and those principles are needed (and will transfer to) the building task. Although unguided or minimally guided instructional approaches are very popular and intuitively appealing, there is evidence from empirical studies that indicate that minimally guided instruction is less effective and less efficient than instructional approaches that place a strong emphasis on guidance of the student learning process [Kirschner et al. 2006]. Thus, by this theory, the Guided Discovery (G) condition show better learning than Exploration (E) and Combined (C) conditions on both the paper pre/post tests and the transfer tower building activity. According to this view, principles of balance are better acquired in G condition, and are needed for effective tower building. In line with this argument, there is an assumption that these acquired principles transfer easily to tower building. There is evidence that guided-discovery learning transfers to tower building from my previous experiments reported above [Yannier, Koedinger, et al. 2015].

However there is still room for improvement. Also, in this view, G and C conditions may have greater enjoyment compared to the E condition, as in the Exploration condition children may have frustration at failure since they don't get any guidance.

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

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

| | Paper Pre/post tests | Tower building | Enjoyment |
|---|---|---|---|
| **You learn what you practice** | G>C>E | E>C>G | |
| **Guidance is better for fundamental principle learning** | G>C>E | G>C>E | G≥C≥E |
| **Exploration is better for learning** | E>C>G | E>C>G | E≥C≥G |
| **Complementary benefits** | C≥G>E | C>G≥E | C≥E≥G |

**Table 3.** Hypotheses based on theories or learning principles.

### 8.4.6 Findings

To see the effects of guided discovery and exploration on learning and engagement, I analyzed our paper pre and post-tests, tower pre and post-tests, and surveys that were given after the game.

A MANOVA analysis was ran to see if there is an overall difference between conditions taking into account the paper pre and post-tests (including results of prediction and explanation items) as well as the tower pre and post-tests. First, I ran a MANOVA analysis for all 3 conditions (combined, exploration and guided). The predictor was taken as the condition variable, and the outcome variables were prediction (pre to post difference), explanation (pre to post difference) and tower scores. This analysis revealed that there was an overall significant effect among the three conditions ($F(2,73)=4.08$, $p=0.0037$).

Secondly, I ran a MANOVA analysis for two conditions at a time, again based on the paper pre and post-tests (including results of prediction and explanation items) as well as the tower pre and post-tests. The MANOVA test for only the combined versus guided conditions showed that there was a trend based on all measures ($F(2,48)=2.09$, $p=0.13$). On the other hand, the MANOVA analysis for the guided versus exploration conditions revealed that there was a significant difference based on all the measures of principle tests and tower scores ($F(2,46)=5.46$, $p<0.01$). Similarly, the results for the MANOVA analysis for the combined versus exploration conditions showed a significant effect ($F(2,47)=4.85$, $p=0.01$). These results may be summarized as C≥G>E (where "≥" suggests a trend and ">" shows a significant difference).

In order to be able to scrutinize the results in more detail, I performed ANOVA analysis for each of the tests and measure separately, which is explained below.

| | Prediction | | | Explanation | | | Tower Building | Enjoyment |
|---|---|---|---|---|---|---|---|---|
| | Pre | Post | Dif | Pre | Post | Dif | Improvement | Score |
| E | 0.67 (0.02) | 0.69 (0.02) | 0.02 (0.02) | 0.15 (0.04) | 0.21 (0.04) | 0.06 (0.04) | 0.17 (0.46) | 0.91 (0.04) |
| G | 0.65 (0.03) | 0.70 (0.02) | 0.05 (0.03) | 0.09 (0.05) | 0.38 (0.05) | 0.29 (0.05) | 1.16 (0.56) | 0.90 (0.03) |
| C | 0.63 (0.02) | 0.73 (0.02) | 0.10 (0.02) | 0.14 (0.04) | 0.37 (0.05) | 0.23 (0.04) | 2.31 (0.50) | 0.94 (0.03) |

**Table 4.** Mean scores of Prediction & Explanation Items in the Principle Tests, Tower pre/post tests and enjoyment measured by the survey. E=Exploration Condition; G=Guided-discovery Condition; C=Combined Condition

### 8.4.6.1 Principle pre/post tests

I analyzed the results for the pre and post-tests to identify any differences between conditions. An ANOVA analysis of the pre-test data confirmed no differences between the conditions at principle pretest. To test for learning, I ran a 1-way ANCOVA with principle post-test score as the outcome variable, principle pre-test as the covariate, and learning type (exploration vs. guided-discovery vs. combined) as fixed factors.

Looking at combined vs. exploration separately for *overall principal test results* (including prediction and explanation items); I found a significant main effect of learning type, with benefits for the combined condition. Average post-test scores were 55% for the combined condition and 45% for the exploration condition ($F(1,49)=10.78$, $p<0.01$, $d=0.58$). This result shows that the combined condition, where children were exposed to both guided-discovery and exploration activities, learned more than the exploration-only condition (Figure 36). Error bars in all the graphs represent standard error. Similarly, looking at guided-discovery vs. exploration separately for overall principal test results (including prediction and explanation items); I found a significant main effect of learning type, with benefits for the guided-discovery condition. Average post-test scores were 54% for the guided discovery condition and 45% for the exploration condition ($F(1,48)=7.44$, $p<0.01$, $d=0.53$). This result shows that the guided discovery condition, where children were exposed to both guided-discovery activities including predict-observe-explain with interactive feedback, learned more than the exploration-only condition (Figure 36). There was no significant difference between combined and guided-discovery conditions when analyzed separately for overall principal results.

The timing between different conditions was matched, so that there wasn't a big difference between the time on task for children in different conditions (based on the video and log data average time for Combined Condition is approximately 15 minutes, similarly the average time for Guided Condition is 15 minutes and the average time for Exploration Condition is 16 minutes). This time only includes the time kids spent interacting with the game and excludes the pre and post-tests.

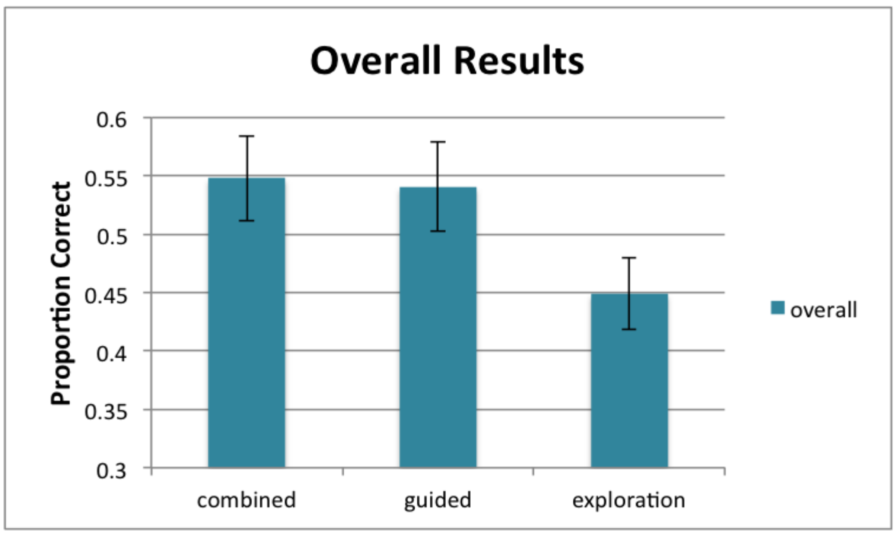



**Figure 36.** Overall post-test results of the paper principle tests including prediction and explanation items. Combined and Guided conditions are significantly better than Exploration (C=G>E).

To investigate if the overall learning benefits for the combined and guided-discovery conditions hold for both prediction items and explanation items, I analyzed each question type separately. To test for learning on the *prediction items*, I ran an ANCOVA with post-test prediction score as the outcome variable, pre-test prediction score as the covariate, and learning type (combined vs. guided-discovery vs. exploration) as fixed factors. The results were slightly different than those from the overall test results. Looking at only prediction items, there was a marginal difference in favor of the combined condition. The average post-test score for the prediction items was 73% for the combined condition, 70% for the guided-discovery condition and 69% for the exploration condition. Analyzing only combined and exploration conditions separately, I ran an ANCOVA with post-test prediction score as the outcome variable, pre-test prediction as the covariate and learning-type (combined vs. exploration) as fixed factors. The results revealed that there was a marginal main effect of learning type ($p=0.06$, $F(1,49)=3.6$, $d=0.38$) with a statistically reliable effect size. Thus children in the combined condition may be better able to predict than the exploration condition. However there was no significant difference for the guided-discovery and exploration conditions when analyzed separately. This result may show that combining the effects of guided-discovery and exploration may lead to better prediction learning outcomes.

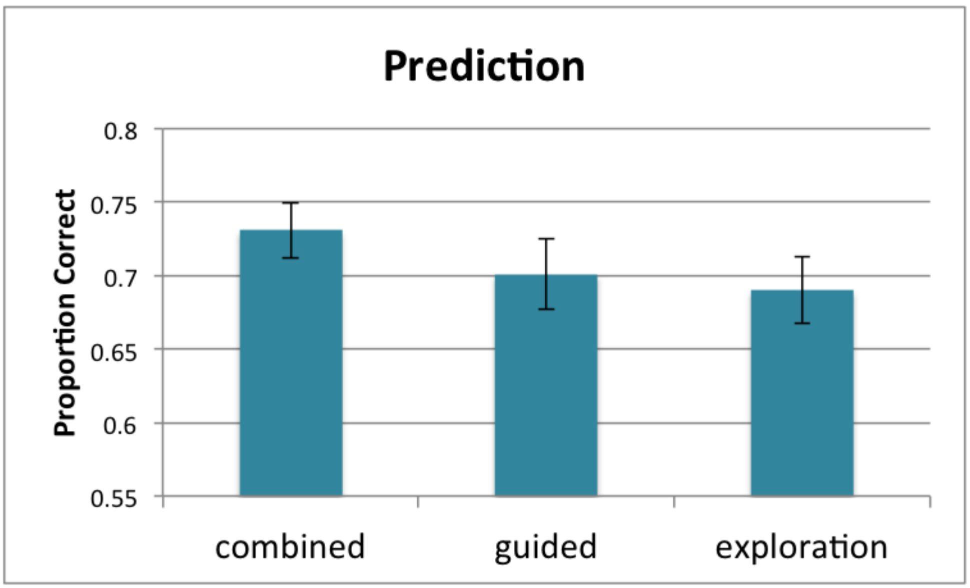



**Figure 37.** Result of prediction items only in the principle post-tests. Guided condition is marginally better than Exploration (C≥E).

Analysis of explanation items on their own led to slightly different results than those for the prediction items. The average post-test score for the explanation items was 36% for the combined condition, 38% for the guided-discovery condition and 20% for the exploration condition. I ran an ANCOVA with post-test explanation score as the outcome variable, pre-test explanation score as the covariate, and learning type (exploration vs. guided-discovery vs. combined) as fixed factors. The results of the ANCOVA test for the combined and exploration conditions, revealed that the combined condition learned to explain the results significantly better compared to the exploration condition ($F(1,49)= 7.72$, $p<0.01$, $d=0.66$). Similarly, analyzing the guided and exploration conditions separately revealed that the combined guided condition learned to explain the results significantly better compared to the exploration condition ($F(1,48)= 9.74$, $p<0.01$, $d=0.75$). However, there was no significant difference between the combined and guided conditions.

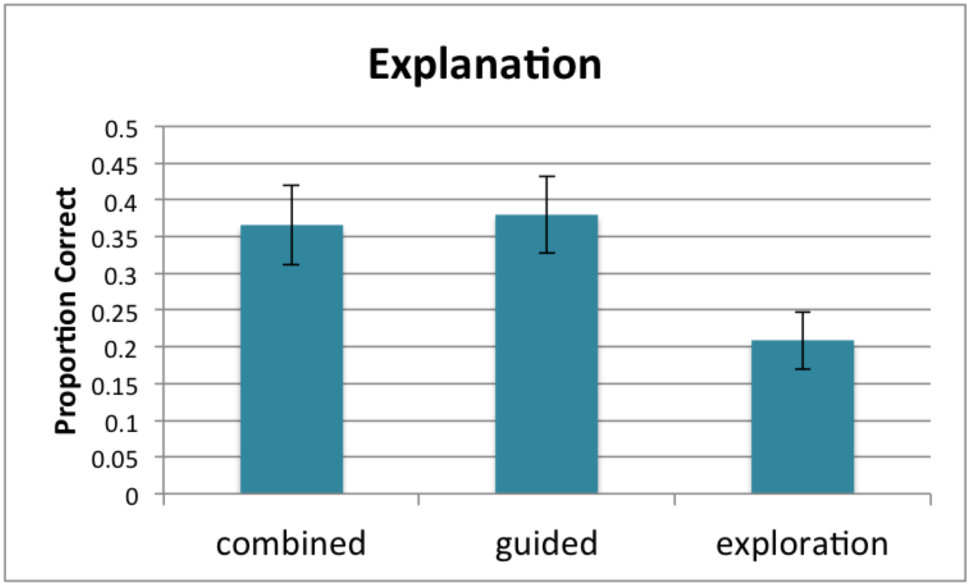



**Figure 38.** Result of explanation items only in the principle post-tests. Both Combined condition and Guided Conditions are significantly better than the Exploration Condition (C=G>E).

Thus even though the exploration condition seems to be able to learn to predict as well as the guided-discovery condition (looking at the results of the prediction items on its own), children in the exploration condition were not able to explain what they learned (as shown in the results of the explanation items). Thus they did not have a deep learning or understanding of the reasons (i.e. why one tower will fall more quickly than another based on the principles of center of mass, symmetry, height, wide-base).

Analyzing the prediction and explanation results, we see that some amount of guidance, present in both the Combined and Guided-discovery conditions enhances explanation. In contrast, some amount of exploration in both the Exploration and Combined conditions does not distinctly enhance prediction. Rather, the combination may enhance prediction.

### 8.4.6.2 Tower-test Results

Tower pre and post-tests were analyzed to see if there was a similar pattern in the results. First of all, pre and post towers that kids built were scored. For the first tower test (where they were asked to build a tower that was more stable than the given tower using the same blocks), we scored each pair's towers according to four principles: height, symmetry, center of mass and width of the base. For each principle, students were given one point if their towers improved from pre- to post-test, -1 for the reverse, and 0 for no change. Comparing pre- and post- towers for the height principle, a shorter post-tower scores 1, a taller post-tower scores -1, and towers of the same height score 0. Likewise, post-towers with more symmetry and a lower center of mass score one for each of those principles. Adding the scores for each principle yielded the student's total score for the first tower test.

For the second tower test (where they were asked to build a tower that would not fall down when the table shook, using the given blocks and using a specific block on the bottom as the base), 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). Again, for each principle, the pair of students was given a 1, 0 or -1 based on the improvement of the tower on these principles. Then the scores for each principle were added to get a total score for this tower test. Then the results of the first and second tests were added together to come up with a total score for the tower tests.

To investigate the transfer of the learning from the game to hands-on real-world activities, I analyzed the overall tower scores. I ran an ANOVA with overall tower score as the outcome variable and learning type (combined vs. guided-discovery vs. exploration) as fixed factors. The results of the ANOVA test for the combined and exploration conditions, revealed that the combined condition transferred significantly better to hands-on experimentation compared to the exploration condition ($F(1,49)= 9.38$, $p=0.0036$, $d=0.92$). When I ran an ANOVA test for the guided and exploration conditions, the results revealed that there was a trend showing guided condition transferred marginally better to hands-on experimentation compared to the exploration condition ($F(1,48)= 2.45$, $p=0.1$, $d=0.40$). Similarly, an ANOVA test for the combined and guided conditions, showed that there was a trend in favor of the combined condition – combined condition transferred marginally better to hands-on experimentation compared to the guided condition ($F(1,50)= 2.01$, $p=0.1$, $d=0.43$). Thus, C≥G≥E and C>E where "≥" represents a marginal difference or trend and ">" represents a significant difference. These results can mean that children who are exposed to both guided and exploration activities can transfer better to real-world hands-on tasks.

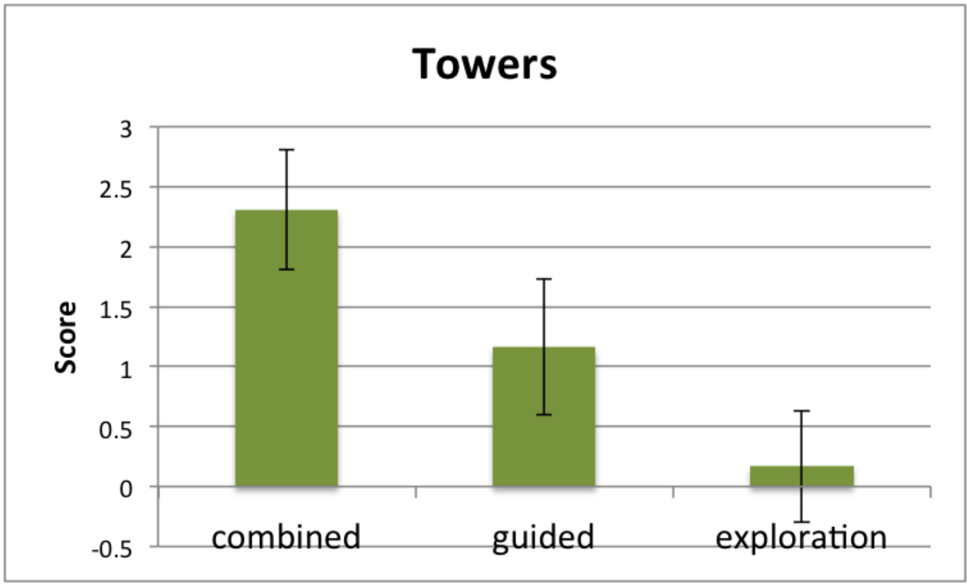



**Figure 39.** Result of tower pre and post-tests. Combined condition is significantly better than Exploration Condition. Guided condition is marginally better than Exploration, and Combined condition is marginally better than Guided condition (C≥G≥E && C>E).

I also analyzed the explanations children made after being asked by the experimenter "Can you explain how you built the tower? Did you have any strategies in mind?" after building the tower in the task they were given. In order to do this, I analyzed and coded the videos of the children for the pre and post tower tests. They got a 1 if they made an explanation that involved the principles from the game (i.e. if they said something that was related to the wide-base, height, symmetry or center of mass principles in the game). For example, if they said: "We tried to make a strong base" or "We made it even", they would get a point for their explanations. In the combined condition, one student explained why they built the tower the way they did by saying: "We learned from the game that the long one has to be on the bottom". On the other hand, some others said they wanted to build a tower that would look like a peacock or museum etc., in which case they did not get any points for their explanations.

I ran an ANCOVA with post-test tower explanation score as the outcome variable, pre-test tower explanation score as the covariate, and learning type (exploration vs. guided-discovery vs. combined) as fixed factors. The results of the ANCOVA test for only the combined and exploration conditions, revealed that the tower explanations of the students in the combined condition were significantly better compared to those in the exploration condition (F(1,49)= 4.04, p=0.05, d=0.45). Similarly, analyzing the guided and exploration conditions separately revealed that the combined guided condition were able to explain the towers they built significantly better compared to the exploration condition (F(1,48)= 3.93, p=0.05, d=0.56). However, there was no significant difference between the combined and guided conditions.

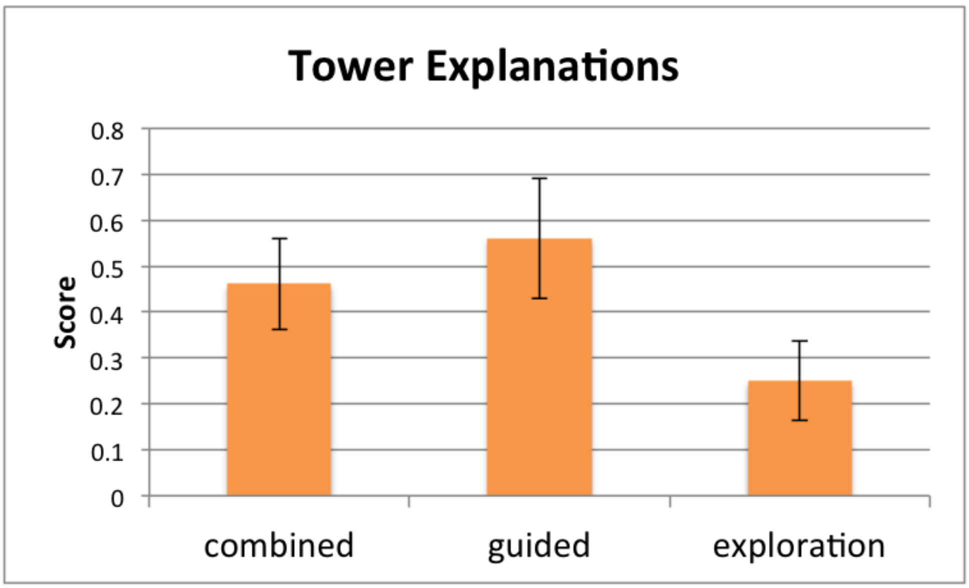



**Figure 40.** Result of tower explanation scores. Both Combined condition and Guided Condition explain the towers they built significantly better than Exploration Condition (C=G>E).

#### 8.4.6.3 Enjoyment

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 5-point likert scale to three different questions: "How much did you like the game?", "Would you like to play it again?", "Would you recommend it to a friend?". 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.

I analyzed the results of the survey (for the sum of the three questions they answered), and an ANOVA showed that there wasn't a significant difference between the three conditions. There seemed to be a ceiling effect as the average of all conditions' scores was above 90%. The average enjoyment score was 94% for the combined condition, 90% for the guided-discovery condition and 91% for the exploration condition. This may mean that children liked the mixed-reality system in general. Some of the kids even made some drawings on the survey questions, putting smileys on the highest scores and writing "I want to play this every day!" as a comment (See Figure 42).


**Figure 41.** Result of enjoyment surveys. There was a ceiling effect and no significant difference between the three conditions.

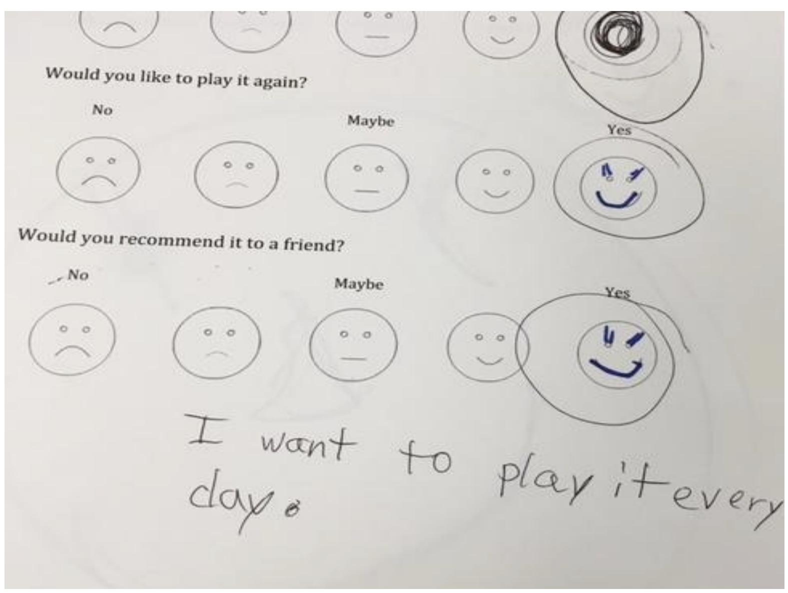



**Figure 42.** Student wrote comments and drew smiley faces on the surveys.

### 8.4.6.4 Log/process data

I also analyzed the log data (the data captured as children interacted with the game during the experiment) to see if there are any patterns in their interaction. Children in the Combined condition interacted with guided discovery activities (including prediction and explanation items) in the sequence of five contrasting cases/tower pairs: "D2-D3", "D1-D2", "B4-B3", "D3-D4", "A2-E2" ("Tasks everyone did" in Table 3) and then they were asked to build a tower in the game that would stay up when the table shakes, then they interacted with "C1-C2" and "A1-A2" after building the tower ("After building" in Table 3). Children in the Guided condition interacted with same seven tower pairs "D2-D3", "D1-D2", "B4-B3", "D3-D4", "A2-E2", "C1-C2", "A1-A2" as in the Combined condition, but note that they did the last two pairs, "C1-C2", "A1-A2", immediately after the others without intervening tower building as in the Combined condition. These pairs thus provide a nice learning process assessment of whether the building done by the combined group leads to any transfer to prediction or explanation. Note, to try to keep overall instructional time similar, the Guided group also interacted with three extra pairs: "D3-B1", "E2-B1", "B1-B2". Table 3 shows the proportion of children in each condition who answered each of the prediction and explanation items correctly for each tower pair.

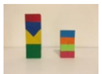
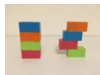
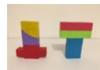
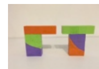
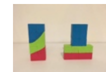
| | Tasks everyone did | | | | | After Building (for Combined) | | Extras (for Guided) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | D2-D3  | D1-D2  | B4-B3  | D3-D4  | A2-E2  | C1-C2  | A1-A2  | D3-B1  | E2-B1  | B1-B2  |
| **Guided** (prediction) | **0.58** | **0.67** | **0.83** | **0.91** | **0.82** | **0.67** | **0.33** | **0.75** | **0.92** | **0.75** |
| **Combined** (prediction) | **0.75** | **0.92** | **1.00** | **1.00** | **0.85** | **0.85** | **0.85** | | | |
| **Guided** (explanation) | **0.66** | **0.75** | **0.50** | **0.75** | **1.00** | **0.50** | **0.58** | **0.92** | **0.75** | **0.92** |
| **Combined** (explanation) | **0.92** | **0.77** | **0.23** | **0.77** | **0.77** | **0.54** | **0.69** | | | |

**Table 3.** Proportion of children that answered each of the prediction and explanation items correctly for each tower pair in the Combined and Guided conditions.

I averaged the proportions of participants in each group (Tasks everyone did, After building and Extras) to see the performance of each group separately overall for different conditions (Shown in Table 4 below). An interesting pattern is apparent. The proportion of children that answer the prediction items correctly after building towers in the Combined condition, seem to be higher than for those who answer the same items with no building task in the Guided condition (0.85 in Combined vs. 0.50 in Guided). On the other hand, the proportion of children who answer the explanation items correctly for the same questions after building towers in the Combined condition is not so different than those in the Guided Condition (0.62 for Combined vs. 0.54 for Guided). This apparent interaction appears to match the result of the prediction items and tower building tasks in the post tests explained above, where children in the Combined condition do better in predicting and transferring to a tower building task, but don't do any better on explanation items.

| | Tasks everyone did | After Building (for Combined) | Extras (for Guided) |
|---|---|---|---|
| **Guided** (prediction) | **0.76** | **0.50** | **0.81** |
| **Combined** (prediction) | **0.90** | **0.85** | |
| **Guided** (explanation) | **0.73** | **0.54** | **0.86** |
| **Combined** (explanation) | **0.69** | **0.62** | |

**Table 4.** Average of the proportion of children that answered the prediction and explanation items correctly in the Combined and Guided conditions.

To test for the statistical reliability of the apparent patterns described above, I performed an ANCOVA analysis with log-data score of after building prediction items as an outcome, prior prediction ability (a combined score of pretest prediction items on paper tests and log-data score of prediction tasks everyone did) as the covariate and learning type (guided-discovery vs. combined) as a fixed factors The ANCOVA test revealed that the students in the Combined condition performed significantly better on the after-building prediction tasks than those in the Guided-discovery condition ($F(1,50)= 9.74$, $p<0.003$, $d=1.04$) (C>G). I ran a similar ANCOVA analysis for the explanation items - this time taking log-data score of after building explanation items as an outcome, prior explanation ability (combination score of pretest explanation items on paper tests and log-data score of explanation tasks everyone did) as the covariate and learning type (guided-discovery vs. combined) as a fixed factor. The ANCOVA revealed no significance difference between the performance of students in the Combined Condition compared to those in the Guided-discovery Condition on the after-building explanation tasks ($F(1,49)= 0.47$, $p=0.50$, $d=0.19$) (C=G).

This interesting process result provides further evidence for the hypothesis that combining guided-discovery and exploration activities may improve learning of prediction but not of explanation. In other words, adding some engineering activities to scientific thinking practice may improve learning outcomes on engineering related measures (such as prediction) without any sacrifice to learning outcomes on science related measures (such as explanation).

#### 8.4.6.5 Qualitative data

Qualitative anecdotes illustrate the students' enjoyment and engagement. Informal review of the video data suggests that many children in the mixed-reality condition were highly engaged. For example, kids said thing such as: "I want to stay here forever", "I want to do this every day for the rest of my life", "Can I come here every day?", "Can we come here every week when we get bored?", "I feel like I could play this forever", "I want to live here!", "Why would you want to leave this place?", "Can we come here again today? Do we have to wait until next week?", One of the children said while playing: "This is the best day of my life! …And I have tennis today, which I barely like…"

They were given a survey to see how much they enjoyed the game after interacting with the game. They were also asked how much they liked the game after they interacted with the game. One of the kids said: "I liked it really really much! I don't know how to explain it! Too good!", while many others said "It was really really fun!", "I loved it. Who wouldn't?", "It was fun! I don't want to stop playing it.". Many of them mentioned that they especially liked building their own towers and testing them on the earthquake table, while some others said they liked making a guess about what would happen and seeing the results. After they were done with the experiment, they wanted to play more, build more towers and test them on the earthquake table.

The fact that children enjoyed the game very much in all conditions (as shown in our survey data) is interesting. Our qualitative data supports these findings in that in all conditions children seemed to enjoy the mixed-reality environment, observing physical phenomena with interactive feedback and guidance from the gorilla character. Children were excited to see if their prediction was right or wrong while interacting with the game. Many of them said "Yes!" and jumped or made a hand gesture when they saw that their prediction about which tower would fall was correct. The gorilla character in the game also seemed to be a motivation and reward for them when they got the correct answer. When the gorilla character started dancing after they predicted the correct tower that was going to fall, some of them started imitating the gorilla, dancing along with it. Many of them also pointed out that they really liked building their own towers and testing them on the earthquake table. They seemed to be very much engaged while building their own towers. One of the girls, who was at first struggling while answering the questions in the guided-discovery tasks, got very much engaged while building her own towers after the game and said: "I'm gonna be a builder when I grow up, because a lot of these didn't fall!" Many children commented after the game that they really enjoyed interacting with the game and would like to come back again. They also started going to the principal and asking if they could do this activity again. When asked what they liked about it, one of them said: "Everything", while another said "My favorite thing is the game!". So, it looks like they enjoyed the system in general.

There were also some signs of learning from the game in the qualitative data, which may help explain the dynamics going on within the game in the different conditions. It seemed like in the Guided and Combined conditions, the explanation menu prompted kids to try formulate an explanation even if they didn't have an idea about what the reason might be and helped them learn the physics principles and reasons behind the physical phenomena. For example, in once case, the pair in the combined condition made a prediction, but did not know the reason. When they saw the explanation menu, the girl asked the boy: "What would we choose in that menu?" Then they found the correct explanation. Afterwards, while they were asked to build a tower in the game they said: "Let's make it so it has a wider base!" They seemed to be strategizing more while building their towers in the combined condition after being guided through the game, than in the exploration condition. For example, after going through the guided-discovery activities with the contrasting cases, when they were asked to build a tower, one of the pairs (in the pilots) built a tower with a wide base. When they tested the tower, they were very happy to see their tower stayed up. One of the girls said: "That's why I put that at the bottom… cos I learned today that if it's like that at the bottom it won't fall!" and then said to her dad "Dad, you should be watching your children!" with pride that she had succeeded. Similarly, another boy built a tower that has a wide base using the small blocks. When he saw that the tower stayed up, he was very excited and started jumping up and down and dancing, saying: "That's because I made a bigger base!!". Then he made another tower this time with other blocks, again having a wide base. When his tower stayed up he said: "All of my ideas are working! I have the bigger base!" (See Figure 43). We did not observe any such comments in the exploration condition when children did not go through the guided-discovery activities. Without guidance from the system, many of the children got fixated and seemed to be making random towers that had interesting shapes such as animal shapes. They said they made a peacock, or a balcony with stairs to go up. One of them said: "We made great wall of China". Thus, this may suggest that children are seeking organizational principles so that they're not simply making choices at random. Offering them learning-productive principles to adopt is actually in service to them when completing an open-ended task. With the help of self-explanation menu and applying the principles they have learned, they can perform an open-ended task in a more structured way as explained above, instead of making random shapes reminding them of animals or other objects they like.

There were also some signs of learning while children were answering the paper post-tests after interacting with the game. One of the students in the Combined condition said: "After playing the game I can answer these questions more easily." while answering the questions in the post-test. Some of them also seemed to pick up words from the game, even if they didn't seem to pay attention during the game. One of them said: "How do you spell symmetrical?" while answering an explanation item in the post-test questions that asked why a symmetrical tower had stayed up. Also, while they were building towers in the post-test some of them indicated that they were using principles they had learned from the game: "We learned from the game that the longer blocks should be on the bottom." These evidences show that children were actually aware of what they learned from the game and could apply it to transfer tasks outside the game.

**Figure 43.** Children were excited while testing their own towers in the game after going through the guided-discovery activities. "That's because I made a bigger base!!", "All of my ideas are working! I have the bigger base!" one of the kids yelled.

A reporter came to observe the experiment one day during the experiment. She did not interact with the children at all during the experiment, however wanted to ask a few questions after they were done with the experiment. "You're a second grade right? Is this the way you learn with your teacher or is this different?" The girl replied: "This is way different! We never do this!", while the other boy said "This is extremely funner than school!". The girl added: "This is probably funner than Kahut!". We figured out that Kahut is a popular game among the children in the school. This was interesting, as it shows that this type of activity is not something that they do in their everyday school experience and even though they play games in school and at home, this system was still different and interesting for them. The reporter asked another pair of students in the combined condition: "Do you think you learned anything?" One of the girls replied: "Yeah! I think we learned that if something's really tiny at the bottom and then something is really heavy at the top (showing with gestures with her hands the shape of the tower), it won't stay!" The reporter asked: "Do you think you knew that before? Did you know that yesterday?" Both of the kids shook their heads saying no. This again shows that children were aware of what they were learning and that they had a deep understanding of the principles they learned. They weren't only able to answer questions on a post-test better, but they could actually consciously articulate about what they had learned.

The reporter also commented that she really liked and was surprised by how the kids collaborated. She said even if they didn't seem to want to collaborate in the beginning, after interacting with the game they started discussing together and collaborating to build towers together. Many of the pairs helped each other while making predictions and explaining during the game. For example, in one of the pairs when the boy was going to click the SHAKE button without discussing, the girl said: "No, don't shake. Discuss!". Then started explaining that the one on the left (D4 in Figure 23) will fall because it's unstable, while the boy said he wasn't sure as both could fall down.

One of the participants started explaining how nobody listens to the teacher in their science and math classes. She said it is really hard for people to learn concepts like measurement, time and length. She suggested that it would be good to expand the game to teach these concepts as well.

Even though children seemed to enjoy the game a lot in general, there were a few things that caused some problems. For example, some of the kids complained about the gorilla talking too much and saying the same thing repetitively especially during the hands-on exploration condition. They said "Shush" while he kept saying: "Can you make a tower that will stay up when the table shakes?". So, it may be good to add more challenges and other scenarios to the hands-on exploration condition that will make it less repetitive. Many of the kids also asked if they could make two different towers and compete with each other during the exploration condition. It may be good to integrate some competition activities as well into the game. There is also research that shows that competition is enjoyable when it's a means to perfect one's skills [Hays 2005]. So, children may get the opportunity to applying knowledge they have learned from the guided-discovery activities and perfecting their skills while competing. They may also get the opportunity to compare and experiment with towers they want to test against each other (e.g. trying to make a tower that is better than a prebuilt tower).

**Figure 44.** Children love building their own towers and testing them on the earthquake table. They don't want to leave even though time is up.

### 8.4.6.6 Discussion

Integrating across the results, an interesting pattern emerges. In the Guided discovery and Combined conditions children receive guidance, through prompts and feedback on their actions, toward enacting a scientific inquiry process (a predict-observe-explain cycle). These conditions yield better explanation and reasoning as measured by their explanation quality both in the principle paper tests and in the tower-building tasks. Thus, having guided-discovery in a mixed-reality environment helps children formulate better, more scientific theories of the physical phenomena they observing. This finding is also supported by the qualitative evidence that children use phrases from the self-explanation prompts when formulating explanations as they build their own towers or answer questions in the post-test.

However, other outcomes indicate how pure guidance may be enhanced by the addition of some exploration. In particular, the results of the prediction items in the paper tests and tower building tasks suggest that children are able to activate explanatory theory in action better when the guided discovery activities are combined with exploratory activities in the mixed-reality system. Children in Combined condition tend to do better than the guided-discovery and exploratory conditions in predicting and transferring to a construction task.

Adding exploration to guided-discovery activities, not only fosters better learning of the balance/physics principles, but also better application of those principles in a hands-on, constructive problem-solving task. Guidance and structure within a mixed-reality setting facilitates a deeper understanding of the principles and helps children formulate better scientific explanations, but the explanatory theory is activated and turned into practice better when combined with exploration activities within the mixed-reality setting. Again, this finding is supported by qualitative evidence, where children point out that they're using knowledge they have built from the game while constructing their own towers within the game in the combined condition (e.g. "That's why I put that at the bottom… cos I learned today that if it's like that at the bottom it won't fall!").

I also compared the results of the experiment with the different theoretical explanations and hypotheses (shown in Table 5 below). Cells in the Table with a "-" indicate evidence from the experiment against the hypothesis. A "+" indicates the experimental results provide solid evidence supporting the hypothesis. A "0" indicates there was a trend in the direction suggested by the hypothesis, but without statistical reliability and/or there is not strong evidence for it or against it.

The "exploration is better for learning" hypothesis suggests that the Exploration condition should do better in both principle tests and tower tasks. If you let kids explore, they should be able to figure out the principles on their own and make better towers. However, our experiment has evidence against this. In our results, Exploration is never better on any measure. The Combined and Guided conditions perform significantly better than the Exploration condition in overall principle test results and, surprisingly, in transferring to hands-on activities in the Tower-Building tests that are directly analogous to Exploration. This result is especially interesting, as the Exploration condition is similar to how what most museum exhibits and many tangibly interfaces today are designed and what many proponents of "constructivism" or discovery learning advocate.

The "You learn what you practice" hypothesis suggests that for the tower building activities, Exploration (E) condition should do better than the Guided discovery (G) and Combined (C) conditions. The experiment had evidence against this hypothesis, as Combined condition performs significantly better than the Exploration condition. Thus, this explanation can is not supported either.

The results are most aligned with and have strongest evidence for the "Complementary benefits" explanation that suggests that having a combination of guided discovery and exploration activities within a mixed-reality setting may be better for learning. Our results for the overall principle tests are aligned with the results in the table: C≥G>E. Similarly, for the tower building tasks, the results from the experiment suggest that the Combined Condition yields better transfer to a hands-on activity significantly than the Exploration condition, and there are trends in favor of the Combined over the Guided Discovery condition and the Guided Discovery over the Exploration conditions. Thus, our results are aligned with the Complementary benefits explanation (C≥G>E).

"Guidance is better for fundamental principle" learning theory suggest that children in the Guided condition should do best for the principle paper and post tests as well as the tower building tasks. We cannot reject this hypothesis completely, as our results suggest that having guidance in the Combined and Guided conditions improves overall learning outcomes compared to the Exploration condition on the principle pre/post tests. However our data also indicates that children in the Combined condition have better outcomes on the Tower building task (i.e. can transfer better to a construction task) which is more aligned with the "Complementary benefits" hypothesis.

Our results for the enjoyment survey do not have any strong evidence for or against the explanations, since there is a ceiling effect. It suggests that children like interacting with our mixed-reality system in general.

Adding some scientific thinking practice to engineering (provided by the guidance within the Combined condition) enhances engineering outcomes (Tower Building and perhaps Prediction). Adding some engineering to scientific thinking practice does not seem to enhance scientific outcomes (Explanation and Tower Explanation). However, children in the Combined condition do get the same scientific outcomes as those in the Guided condition, despite spending less time on scientific thinking/guided discovery activities (7 guided-discovery activities in C versus 10 in G condition).

Thus it looks like combining the guided-discovery and exploration activities may lead to better learning, deeper understanding of principles and better transfer to real-world building activities. The effect of combined activities may also be further enhanced by a more intelligent combination where the system detects the progress and learning of the children and gives personalized exploration or guided-discovery activities accordingly. The positive outcomes from the log data analysis indicate that such dynamic monitoring may be feasible.

| | Paper Pre/post tests | Tower building | Enjoyment |
|---|---|---|---|
| **You learn what you practice** | G>C>E<br>0 | E>C>G<br>- | |
| **Guidance is better for fundamental principle learning** | G>C>E<br>0 | G>C>E<br>0 | G≥C≥E<br>0 |
| **Exploration is better for learning** | E>C>G<br>- | E>C>G<br>- | E≥C≥G<br>0 |
| **Complementary benefits** | C≥G>E<br>+ | C>G≥E<br>+ | C≥E≥G<br>0 |

**Table 5.** Hypotheses matched with the results from the experiment. "+" means that there is evidence from the experiment that supports the hypothesis, "-" indicates that the experiment disproves the hypothesis, "0" indicates that there is not strong evidence from the experiment that proves or disproves the hypothesis.

# Chapter 9. Future Work

### 9.1 Extending to Different Content Areas Based on Teacher Feedback

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

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

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

One future direction is to create new activities and science curriculum that our can be extended to, to create a reusable platform that combines physical and virtual worlds to improve children's science and inquiry learning and enjoyment. This system, NoRILLA (Novel Research-based Intelligent Lifelong Learning Apparatus) could be applied to many different content areas to teach children different concepts. Some possible content areas include Balancing and Weighing, Forces and Motion, Planetary Systems, Density, Weather, Human Body and Simple Machines. For example, one possible future design may be a Balance Scale that is integrated with a game that teaches kids the principles of balance and moment of inertia.

### 9.2 Intelligent Science Exhibits and Stations

We aim to extend this work to create a new genre of *Intelligent Science Exhibits* in museums, science centers and other informal and formal learning spaces, that combine proven intelligent tutoring system approaches with camera-based vision sensing to add a new layer to hands-on museum exhibits. This intelligent layer will provide personalized interactive feedback and scaffolding to museum visitors while they experiment with physical objects in the real world. We plan to extend the research in a museum setting, adding more personalized feedback to investigate if/how Intelligent Science Exhibits can improve children's science and inquiry learning and enjoyment/excitement in a museum setting. We also plan to investigate how Intelligent Science Exhibits can foster collaboration and productive dialogue among families beyond and above the traditional museum exhibits that are mostly based on merely physical exploration. More personalized feedback and challenges can be added to EarthShake: including an adaptive data-driven model enabling the generation of personalized feedback. For example, if a child is having a hard time with the height principle, and is building tall structures that don't withstand the earthquake, he/she can be given more contrasting cases that target teaching these principles. Also, more challenges can be added to make the game more engaging for students: e.g. if they can build a tower that withstands the earthquake, they will be given the challenge of making a tower that is taller than a certain height that will still withstand the earthquake.

# Chapter 10. Conclusions and Contributions

My dissertation consists of two main parts of work. The first part focuses on creating a mixed-reality system combining physical and virtual worlds and investigating how and why a mixed-reality system can improve learning and enjoyment compared to a screen-only alternative (such as a laptop or tablet version). The second part focuses on how a mixed-reality system can be designed to be most effective for learning and enjoyment, investigating whether guided-discovery, exploration or a combination of both would maximize learning based on multiple different outcomes and measures.

With this work, I aim to make several different contributions:

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

Second, I investigate how to create a mixed-reality environment that is optimized for learning and enjoyment. What are the features that lead to more learning in a mixed-reality game (guided-discovery versus exploration/problem solving versus a combination of both)? My work contributes to the literature on exploration and guided-discovery learning, by demonstrating that having guided discovery activities in a mixed-reality setting can improve children's fundamental principle learning by helping them formulate explanations. It also shows that combining an engineering approach with scientific thinking practice can lead to better engineering outcomes such as transferring to constructive hands-on activities in the real world.

Lastly, my work aims to make a contribution from the design perspective by creating a new educational system that bridges physical and virtual environments to improve children's learning and enjoyment in a collaborative way, fostering productive dialogue and scientific curiosity to be used in museum, school and home settings. With an iterative and evolving design approach (Figure 45), my work sheds insight to how a mixed-reality system can be designed to be effective in museum and school settings including design features (e.g. including self explanation menus and interactive feedback that accompanies physical observation) and scenarios (encouraging predict/observe/explain inquiry skills) that can be extended to different content areas and inform early science learning for young children.


**Figure 45.** Iterative evolving design approach was used while building the system.

# 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. *Build. Dialogue Syst. Tutor. Appl. Pap. 2000 AAAI Fall Symp.* (2000), 65–73.

Louis Alfieri, Patricia J. Brooks, Naomi J. Aldrich, and Harriet R. Tenenbaum. 2011. Does discovery-based instruction enhance learning? *J. Educ. Psychol.* 103, 1 (2011), 1.

Alissa N. Antle. 2013. Research opportunities: Embodied child–computer interaction. *Int. J. Child-Computer 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.

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 fifth international conference on Tangible, embedded, and embodied interaction*. 85–92.

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-Volume 1*. 153–160.

Ruth C. Clark and Richard E. Mayer. 2016. *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning*, John Wiley & Sons.

Douglas H. Clements. 1986. Effects of Logo and CAI environments on cognition and creativity. *J. Educ. Psychol.* 78, 4 (1986), 309.

Douglas H. Clements. 1990. Metacomponential development in a Logo programming environment. *J. Educ. Psychol.* 82, 1 (1990), 141.

Susan Wagner Cook, Zachary Mitchell, and Susan Goldin-Meadow. 2008. Gesturing makes learning last. *Cognition* 106, 2 (2008), 1047–1058.

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*

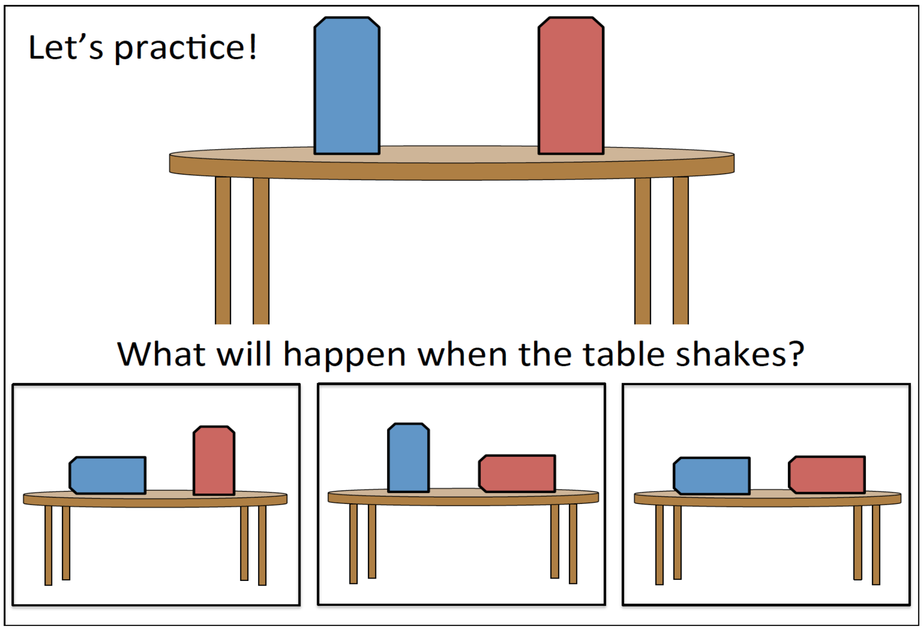
*SIGCHI conference on Human factors in computing systems*. 245–252.
Johannes Engelkamp and Hubert D. Zimmer. 1989. Memory for action events: A new field of research. *Psychol. Res.* 51, 4 (1989), 153–157.
Arthur M. Glenberg. 1997. What memory is for: Creating meaning in the service of action. *Behav. Brain Sci.* 20, 01 (1997), 41–50.
Anuradha A. Gokhale. 1995. Collaborative learning enhances critical thinking. (1995).
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.
Cindy Harvel. 2010. Guided discovery learning. *Faith-Based Educ. That Constr. A Creat. Dialogue between Contructivism Faith-Based Educ. Wipf Stock Publ.* (2010), 169–172.
Harlene Hayne, Jane Herbert, and Gabrielle Simcock. 2003. Imitation from television by 24- and 30-month-olds. *Dev. Sci.* 6, 3 (2003), 254–261.
Robert T. Hays. 2005. *The effectiveness of instructional games: A literature review and discussion*,
Philip H. Henning. 2004. Everyday cognition and situated learning. *Handb. Res. Educ. Commun. Technol.* (2004), 143–168.
Autumn B. Hostetter and Martha W. Alibali. 2008. Visible embodiment: Gestures as simulated action. *Psychon. Bull. Rev.* 15, 3 (2008), 495–514.
Kodi R. Jeffery-Clay. 1998. Constructivism in museums: How museums create meaningful learning environments. *J. Museum Educ.* 23, 1 (1998), 3–7.
Paul A. Kirschner, John Sweller, and Richard E. Clark. 2006. Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educ. Psychol.* 41, 2 (2006), 75–86.
David Klahr and Sharon McCoy Carver. 1988. Cognitive objectives in a LOGO debugging curriculum: Instruction, learning, and transfer. *Cogn. Psychol.* 20, 3 (1988), 362–404.
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, Albert T. Corbett, and Charles Perfetti. 2012. The Knowledge-Learning-Instruction framework: Bridging the science-practice chasm to enhance robust student learning. *Cogn. Sci.* 36, 5 (2012), 757–798.
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.
Alex Kozulin. 2003. *Vygotsky's educational theory in cultural context*, Cambridge University Press.
Jean Lave. 1991. Chapter 4 Situating Learning in Communities of Practice. *Perspect. Soc. Shar. Cogn.* 2 (1991), 63–82. DOI:http://dx.doi.org/10.1037/10096-003
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 et al. 2012. Translating the Statistical Representation of the Effects of Education Interventions into More Readily Interpretable Forms. *Natl. Cent. Spec. Educ. Res.* (2012).
Paul Marshall, Peter C.H. Cheng, and Rosemary Luckin. 2010. Tangibles in the balance: a

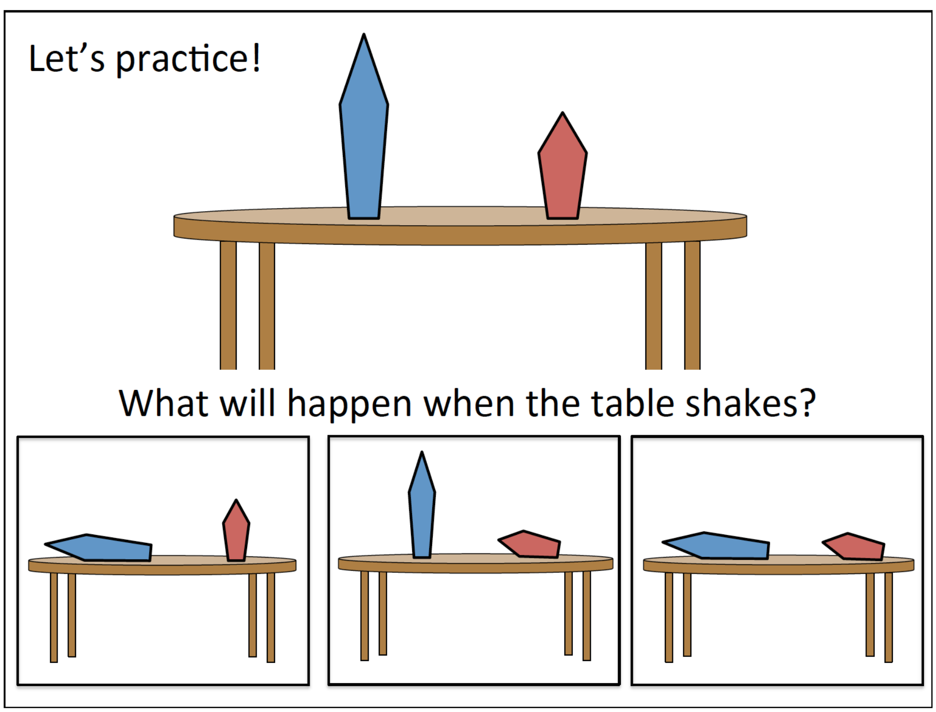
discovery learning task with physical or graphical materials. In *Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction*. 153–160.
Paul Marshall, Yvonne Rogers, and Eva Hornecker. 2007. Are tangible interfaces really any better than other kinds of interfaces ? *CHI'07 Work. Tangible User Interfaces Context Theory* , i (2007), 28.
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. DOI:http://dx.doi.org/10.1207/s15516709cog0000_15
Craig S. Miller, Jill Fain Lehman, and Kenneth R. Koedinger. 1999. Goals and learning in microworlds. *Cogn. Sci.* 23, 3 (1999), 305–336.
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. (2004).
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
Seymour Papert. 1980. *Mindstorms: Children, computers, and powerful ideas*, Basic Books, Inc.
Roy D. Pea and D. Midian Kurland. 1984. On the cognitive effects of learning computer programming. *New Ideas Psychol.* 2, 2 (1984), 137–168.
Laurel Puchner, Robyn Rapoport, and Suzanne Gaskins. 2001. Learning in children's museums: is it really happening? *Curator Museum 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. 1998. Digital manipulatives. *Proc. SIGCHI Conf. Hum. factors Comput. Syst. - CHI '98* (1998), 281–287. DOI:http://dx.doi.org/10.1145/274644.274684
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: Teleoper. Virtual Environ.* 11, 6 (December 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 *CHI '99 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '99. New York, NY, USA: ACM, 272–273. DOI:http://dx.doi.org/10.1145/632716.632883
Leona Schauble, Leopold E. Klopfer, and Kalyani Raghavan. 1991. Students' transition from an engineering model to a science model of experimentation. *J. Res. Sci. Teach.* 28, 9 (1991), 859–882.
Bertrand Schneider, Engin Bumbacher, and Paulo Blikstein. 2015. Discovery versus direct instruction: Learning outcomes of two pedagogical models using tangible interfaces. In *Exploring the material conditions of learning: opportunities and challenges for CSCL," the Proceedings*

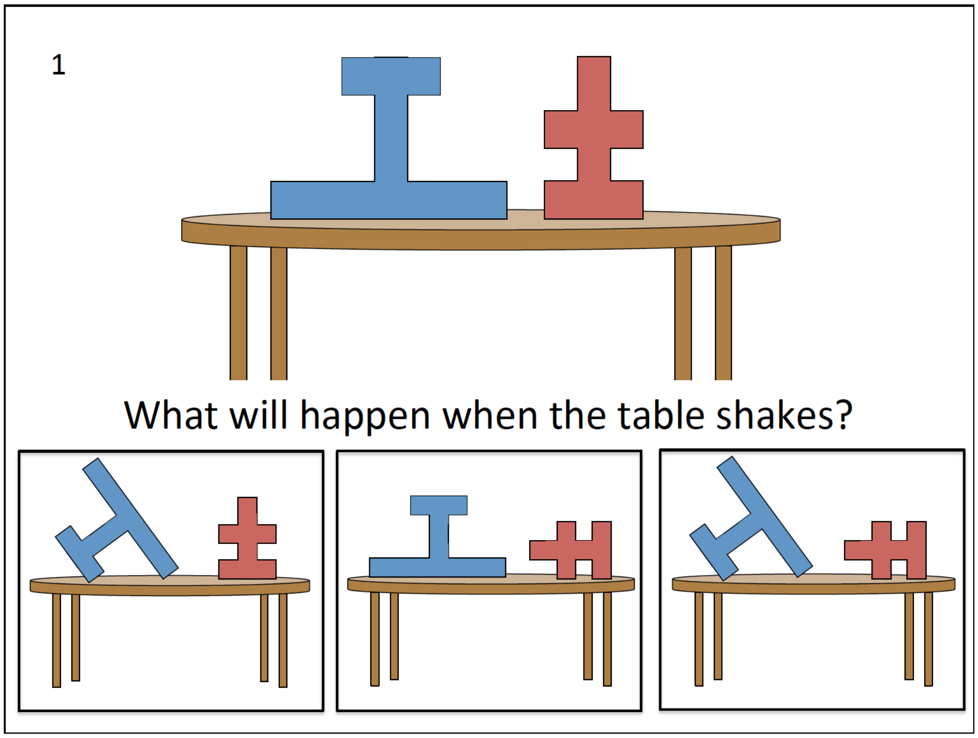
*of the Computer Supported Collaborative Learning (CSCL) Conference*. 364–371.
Bertrand Schneider, Patrick Jermann, Guillaume Zufferey, and Pierre Dillenbourg. 2011. Benefits of a tangible interface for collaborative learning and interaction. *IEEE Trans. Learn. Technol.* 4, 3 (2011), 222–232. DOI:http://dx.doi.org/10.1109/TLT.2010.36
Tia Shelley, Leilah Lyons, Moira Zellner, and Emily Minor. 2011. Evaluating the Embodiment Benefits of a Paper-based Tui for Educational Simulations. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '11. New York, NY, USA: ACM, 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, & Ulrich Hoppe, eds. *Designing for Change in Networked Learning Environments: Proceedings of the International Conference on Computer Support for Collaborative Learning 2003*. Dordrecht: Springer Netherlands, 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. New York, NY, USA: ACM, 118–126. DOI:http://dx.doi.org/10.1145/1999030.1999044
Juhani E. Tuovinen and John Sweller. 1999. A comparison of cognitive load associated with discovery learning and worked examples. *J. Educ. Psychol.* 91, 2 (1999), 334.
Sherry Turkle. 2015. *Reclaiming conversation: The power of talk in a digital age*, Penguin Press HC.
B. Ullmer and H. Ishii. 2000. Emerging Frameworks for Tangible User Interfaces. *IBM Syst. J.* 39, 3-4 (July 2000), 915–931. DOI:http://dx.doi.org/10.1147/sj.393.0915
John Underkoffler and Hiroshi Ishii. 1998. Illuminating Light : An Optical Design Tool with a Luminous-Tangible Interface. *Proc. ACM Conf. Hum. Factors Comput. Syst.* , April (1998), 542–549. DOI:http://dx.doi.org/10.1145/274644.274717
Lev Vygotsky. 1978. Interaction between learning and development. *Readings Dev. Child.* 23, 3 (1978), 34–41.
Erin Walker and Winslow Burleson. 2012. Using need validation to design an intelligent tangible learning environment. *Proc. 2012 ACM Annu. Conf. Ext. Abstr. Hum. Factors Comput. Syst. Ext. Abstr. - CHI EA '12* (2012), 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, Ali Israr, Jill Fain Lehman, and Roberta L. Klatzky. 2015. FeelSleeve: Haptic feedback to enhance early reading. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 1015–1024.
Nesra Yannier, Kenneth R. Koedinger, and Scott E. Hudson. 2016. Apparatus and Methods for Tangible Collaborative Learning. (2016).
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. New York, NY, USA: ACM, 1045–1054. DOI:http://dx.doi.org/10.1145/2702123.2702397

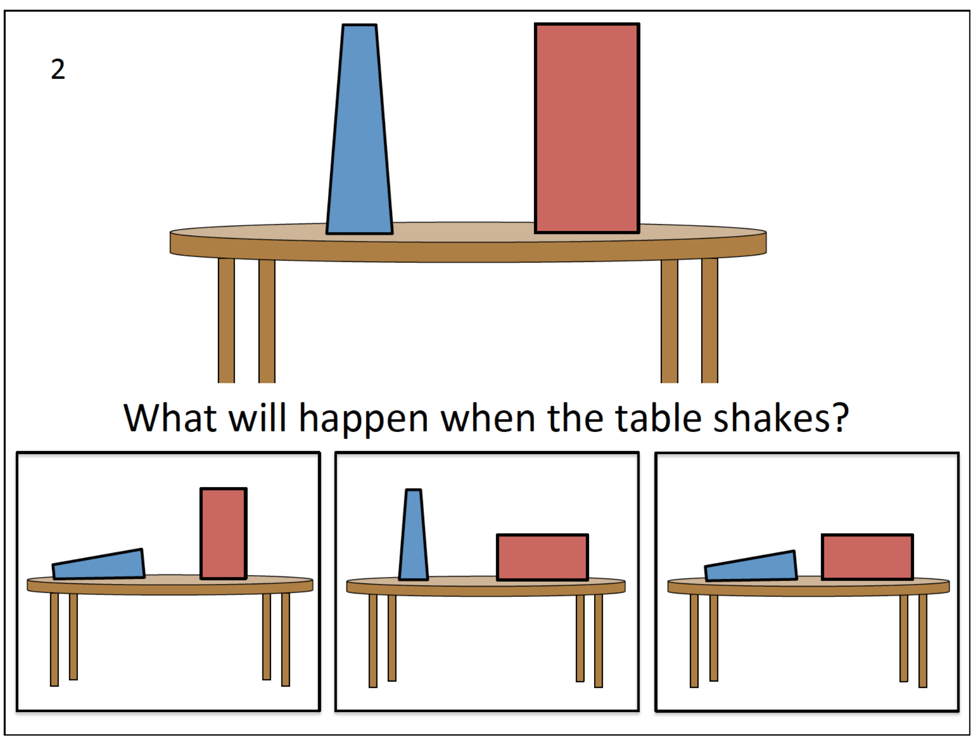
Nesra Yannier, Kenneth R. Koedinger, and Scott E. Hudson. 2013. Tangible Collaborative Learning with a Mixed-Reality Game: EarthShake. In H. Chad Lane, Kalina Yacef, Jack Mostow, & Philip Pavlik, eds. *Artificial Intelligence in Education: 16th International Conference, AIED 2013, Memphis, TN, USA, July 9-13, 2013. Proceedings*. Berlin, Heidelberg: Springer Berlin Heidelberg, 131–140. DOI:http://dx.doi.org/10.1007/978-3-642-39112-5_14

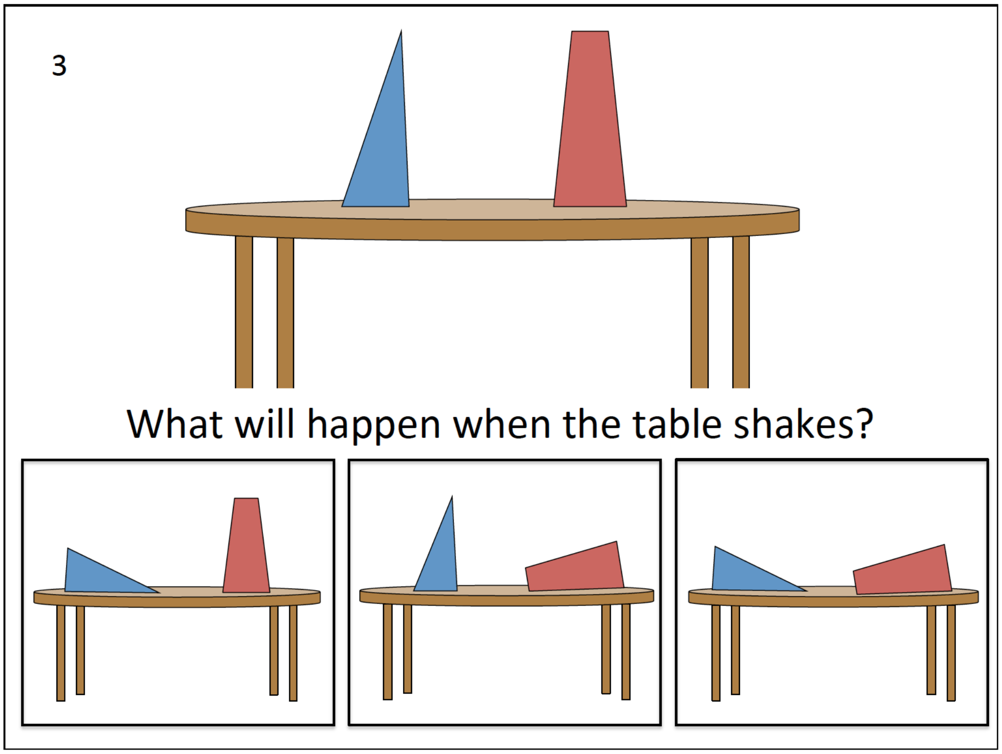
# Appendix I: Pre and Post-Tests used in Experiments – Version A

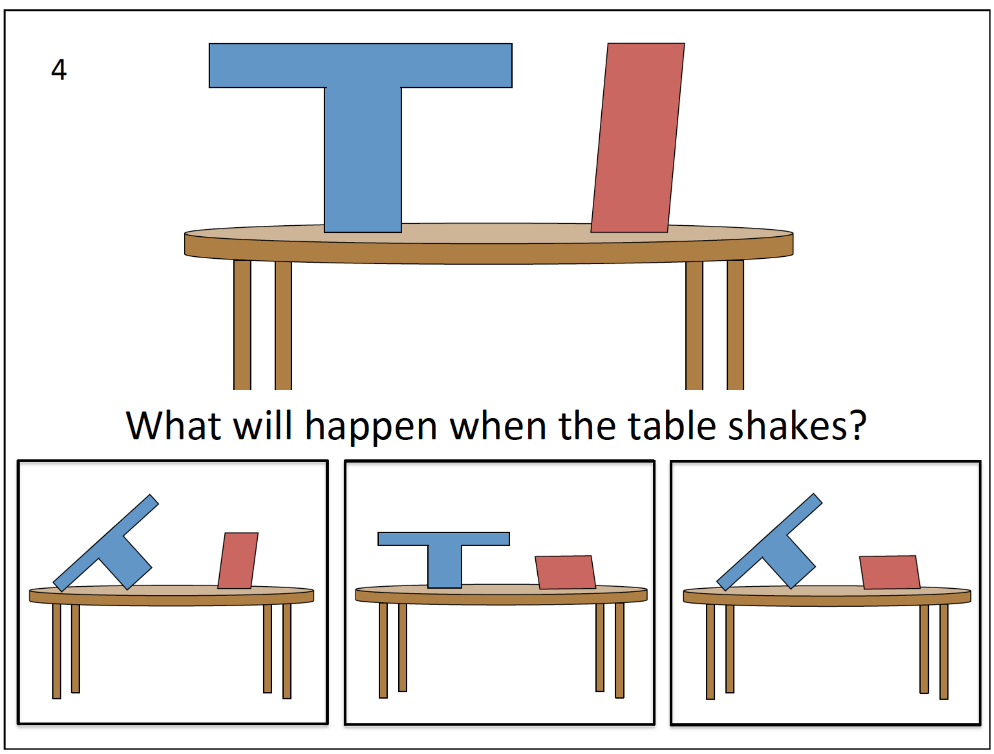


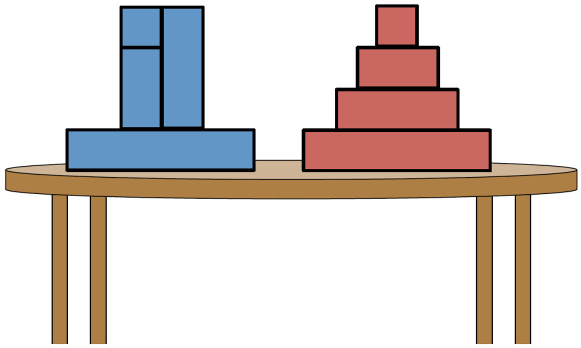

1
What will happen when the table shakes?


2
What will happen when the table shakes?

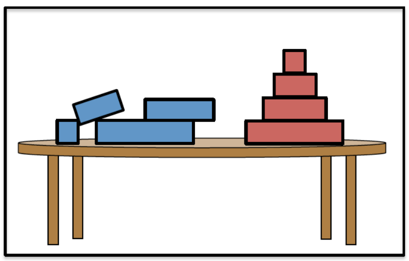

3
What will happen when the table shakes?

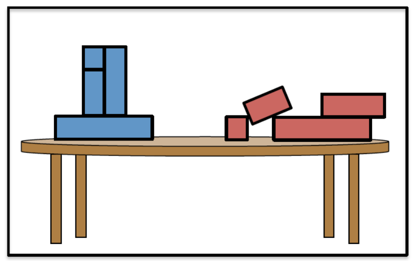

4
What will happen when the table shakes?

5



What will happen when the table shakes?



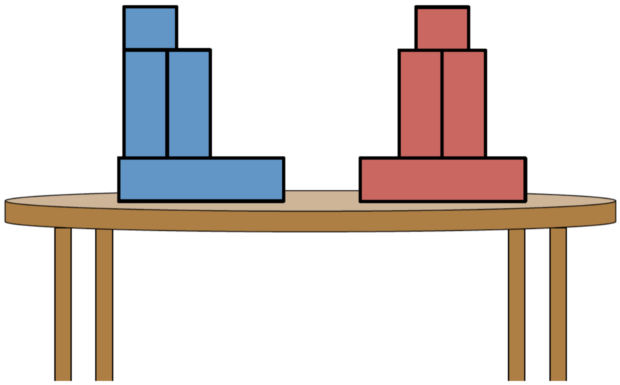


6

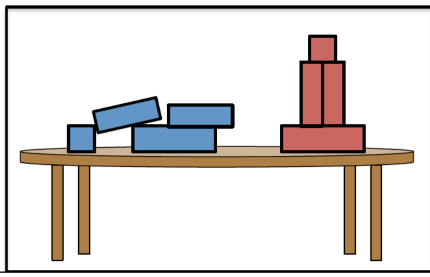
# Why did this one stay up?

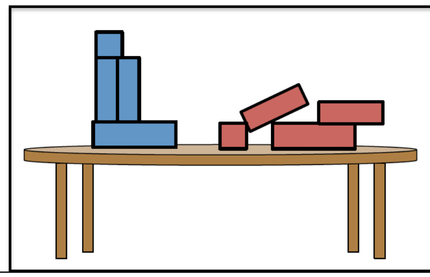
**Because it is ______________________________**

7



What will happen when the table shakes?



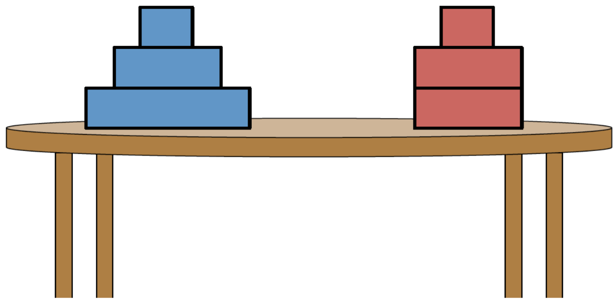


8

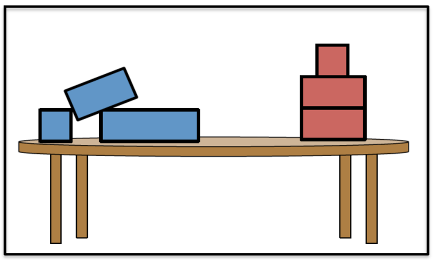
# Why did this one stay up?

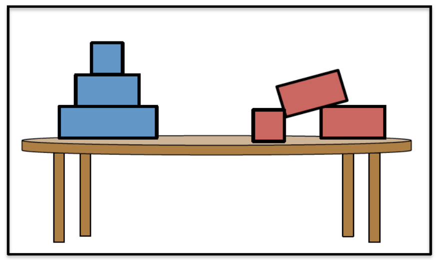
**Because it is ______________________________**

9



What will happen when the table shakes?



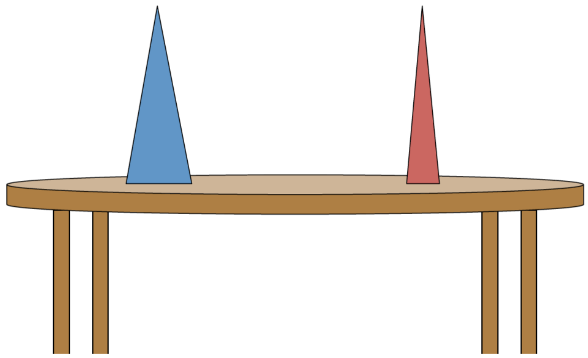


10

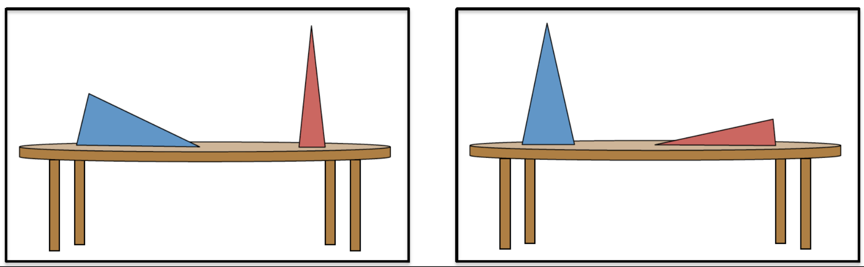
# Why did this one stay up?

**Because it is ______________________________**

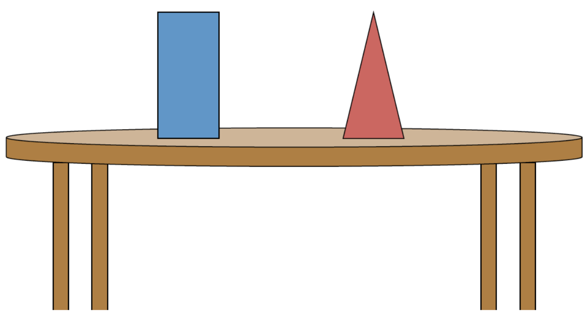
11



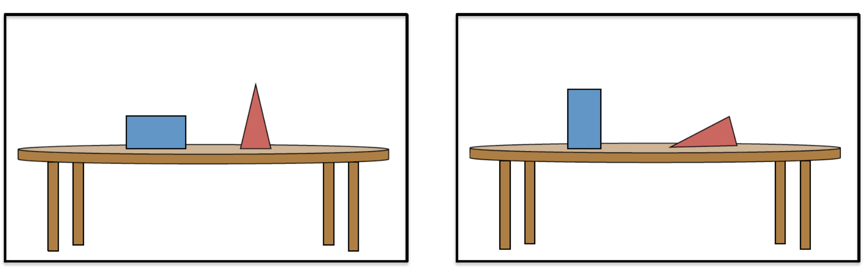
What will happen when the table shakes?



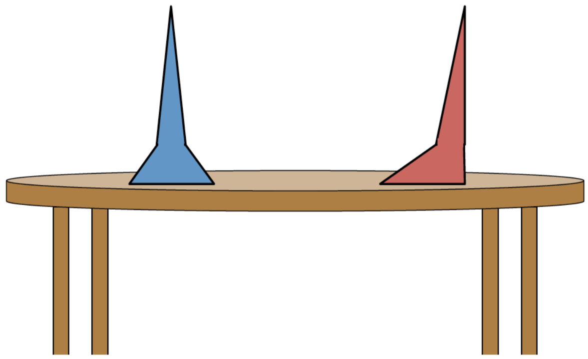
12



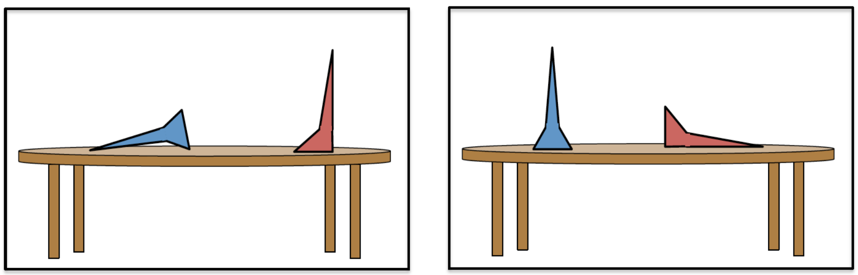
What will happen when the table shakes?



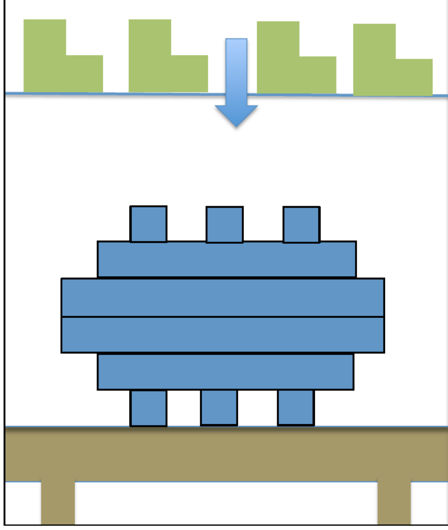
13



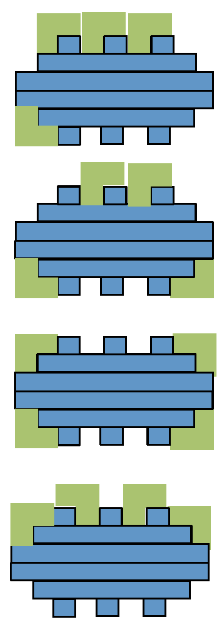
What will happen when the table shakes?

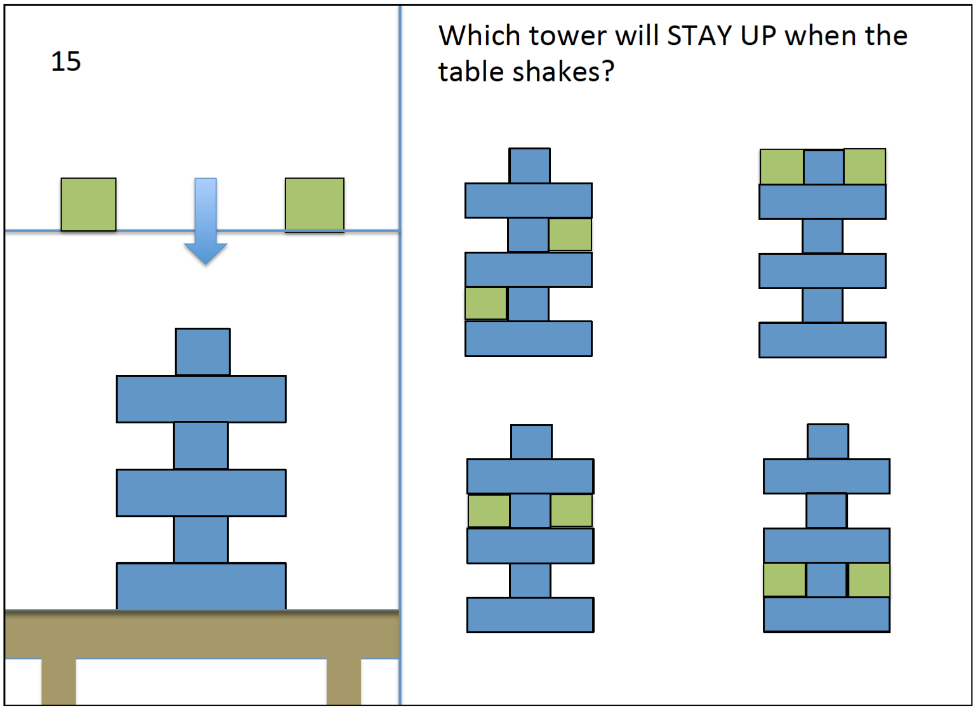


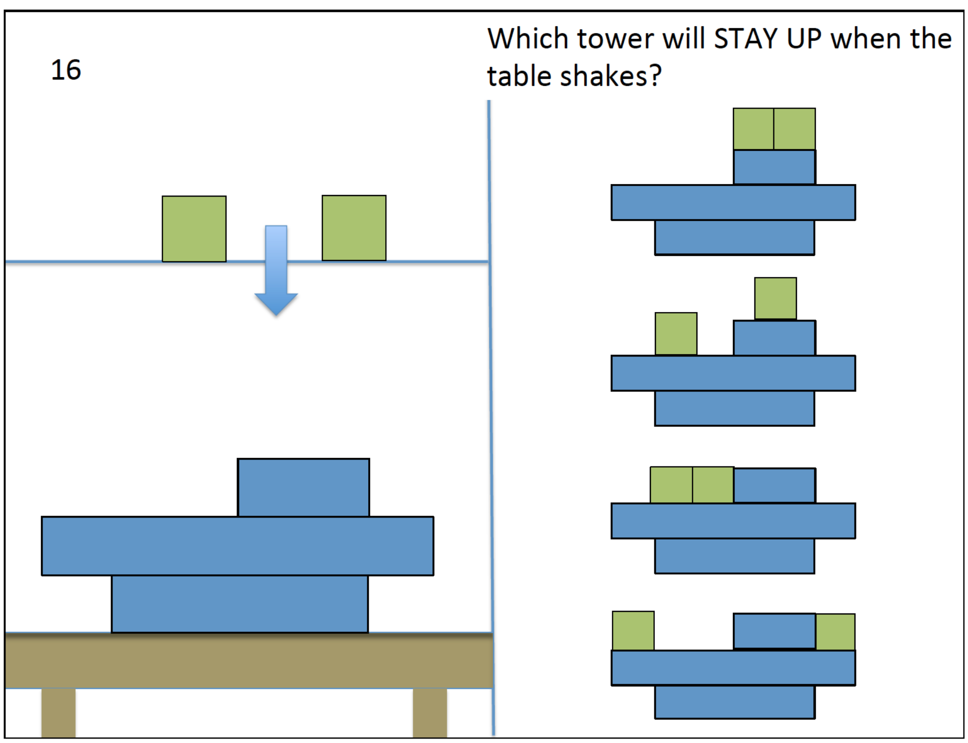
14



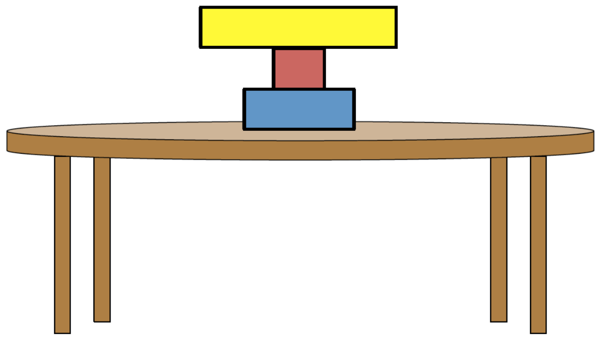
Which tower will STAY UP when the table shakes?




15
Which tower will STAY UP when the table shakes?


16
Which tower will STAY UP when the table shakes?

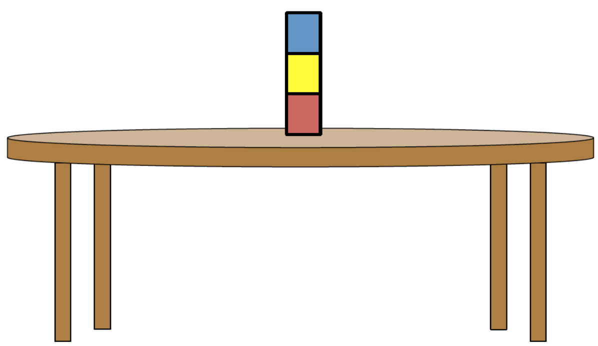
17



This tower will quickly fall when the table shakes.
How could you make it STAY UP? Choose one.

a. Move the blue block to the top
b. Switch the yellow and red blocks
c. Switch the blue and red blocks

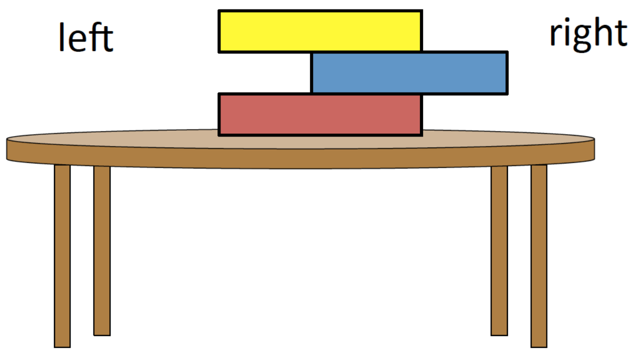
18



This tower will quickly fall when the table shakes.
How could you make it STAY UP? Choose one.

a. Use a skinnier yellow block
b. Use a wider red block
c. Use a skinnier red block

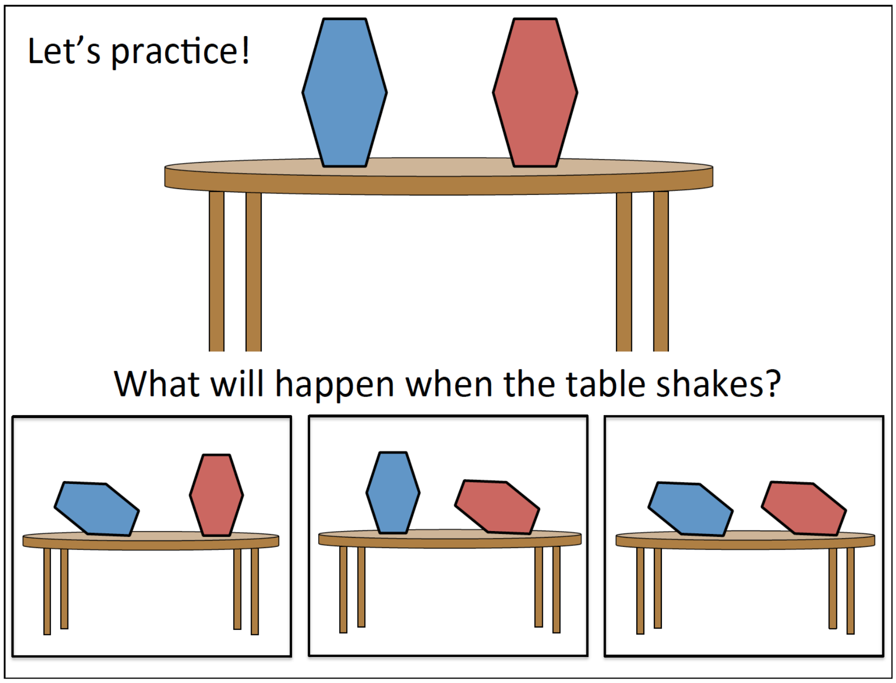
19

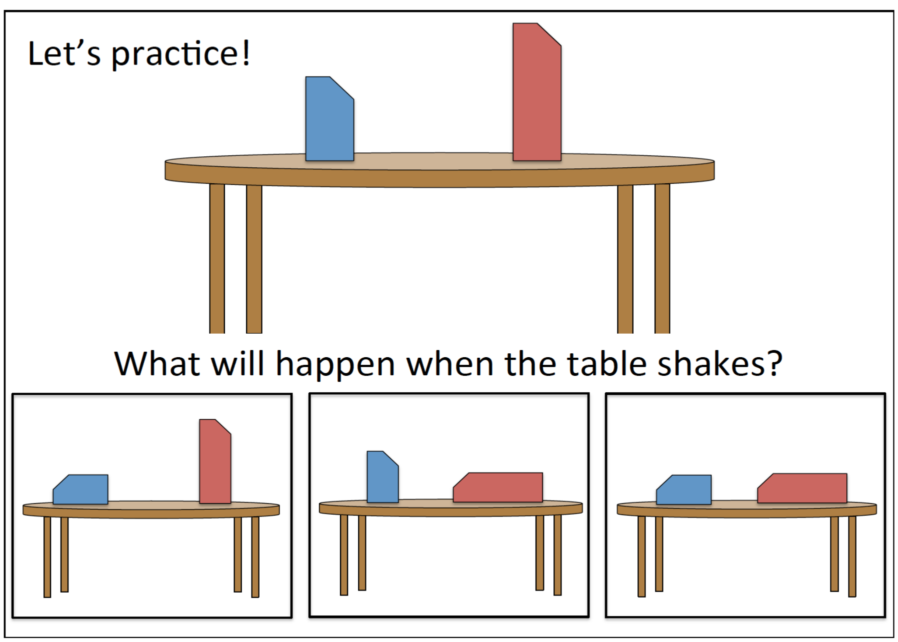


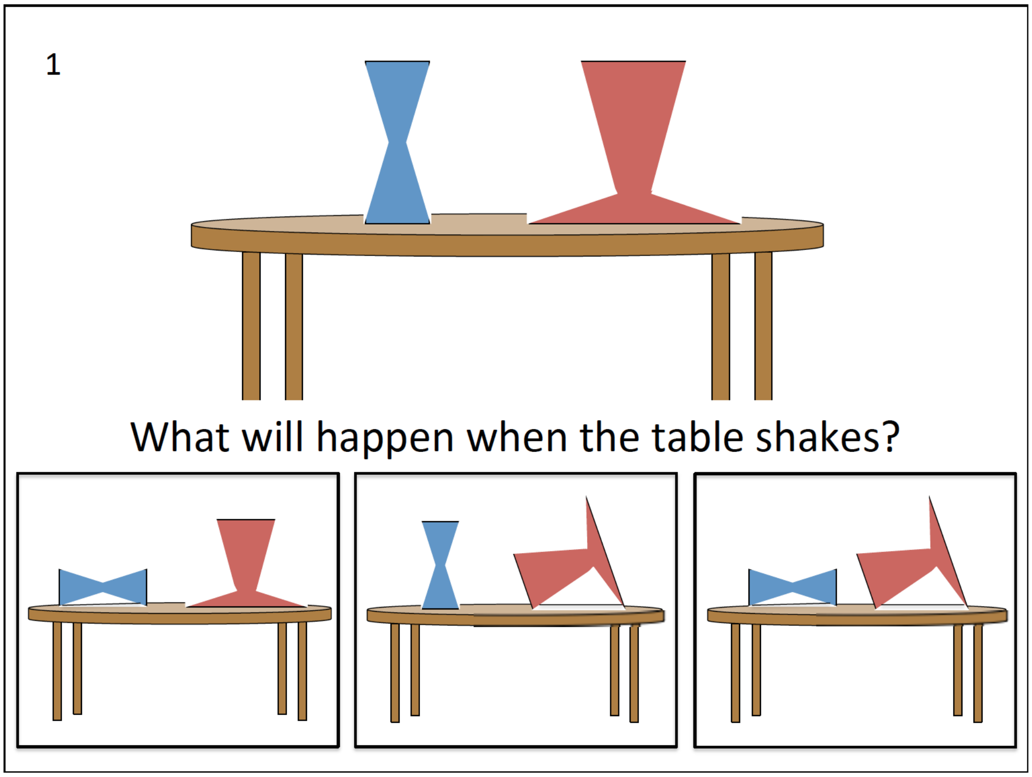

This tower will quickly fall when the table shakes. How could you make it STAY UP? Choose one.

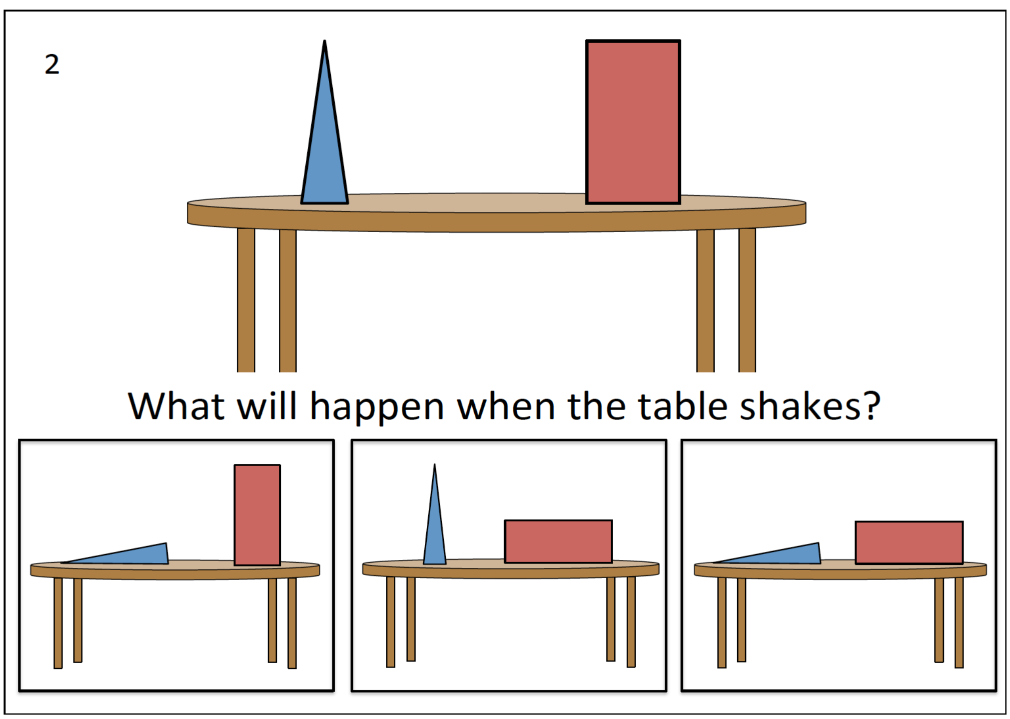
a. Move the yellow block to the right
b. Move the blue block to the left
c. Take the yellow block away

# Appendix II: Pre and Post-Tests used in Experiments – Version B

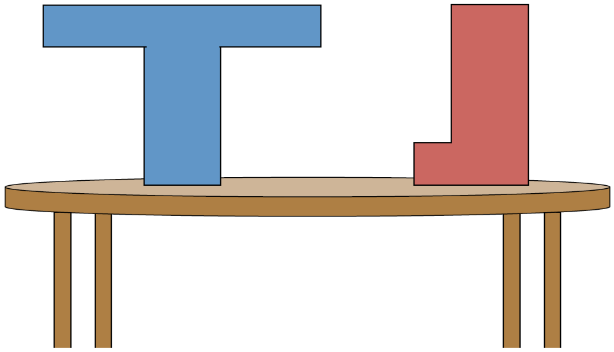


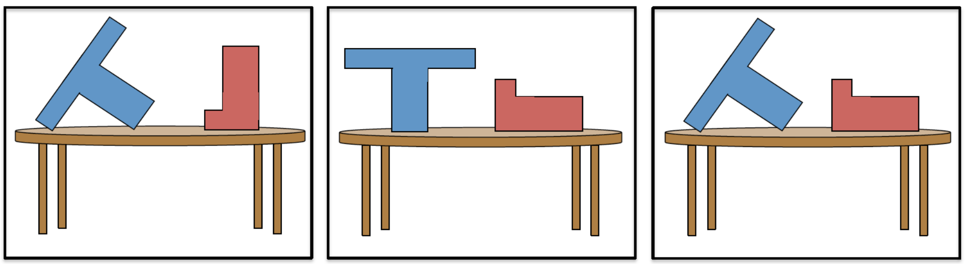

1
What will happen when the table shakes?


2
What will happen when the table shakes?

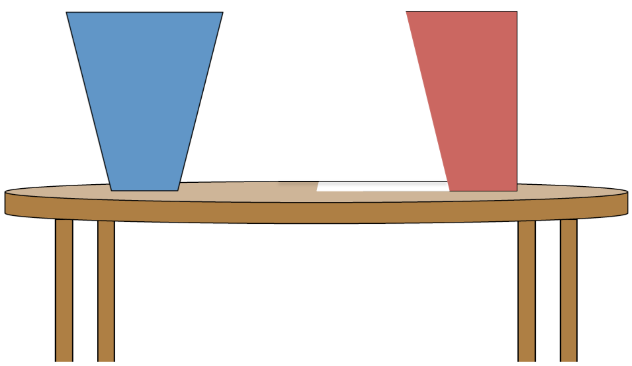
3



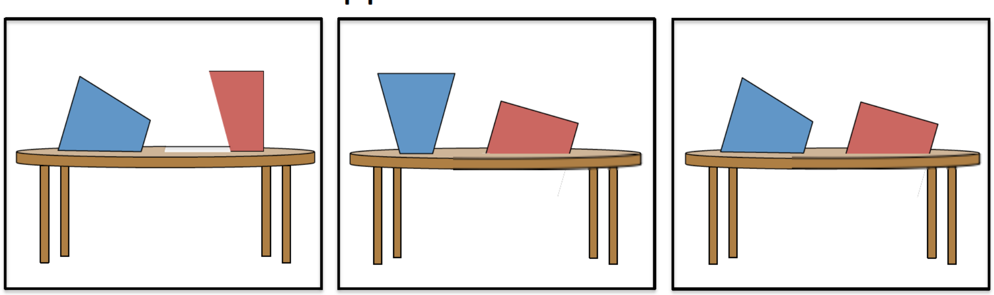
What will happen when the table shakes?



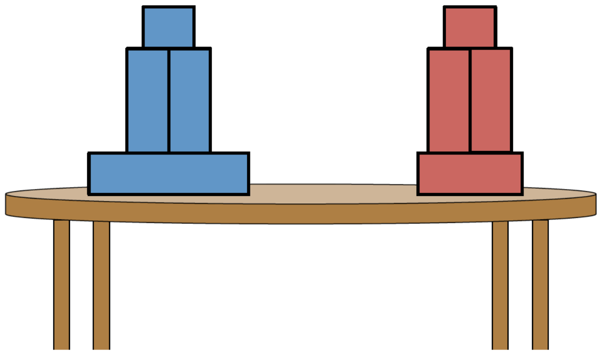
4



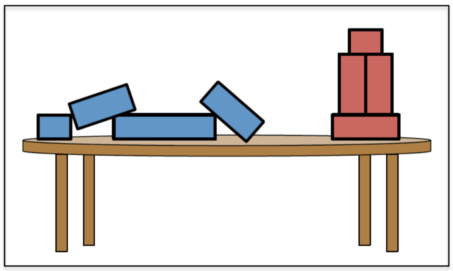
What will happen when the table shakes?

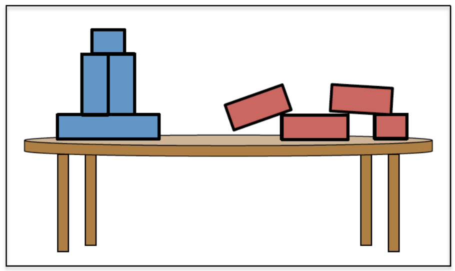


5



What will happen when the table shakes?



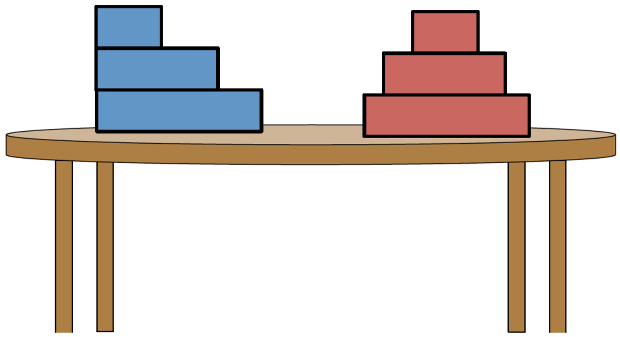


6

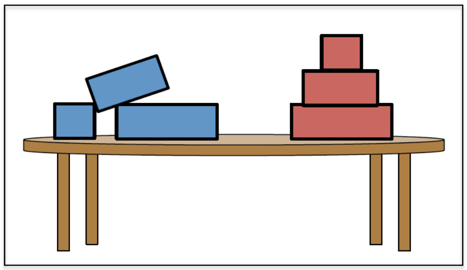
# Why did this one stay up?

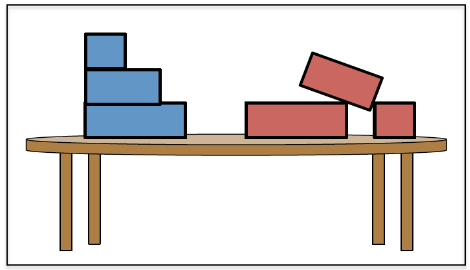
**Because it is ______________________________**

7



What will happen when the table shakes?



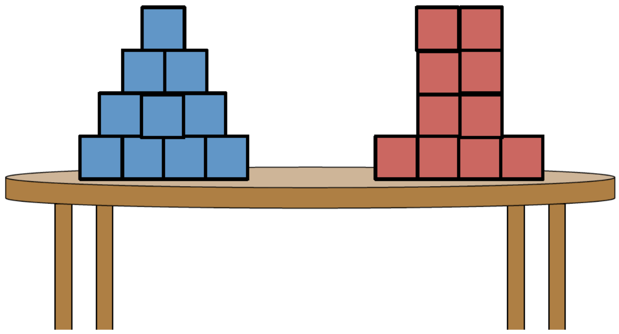


8

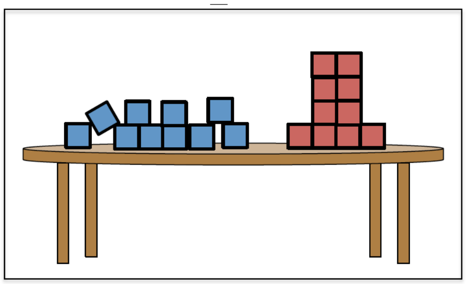
# Why did this one stay up?

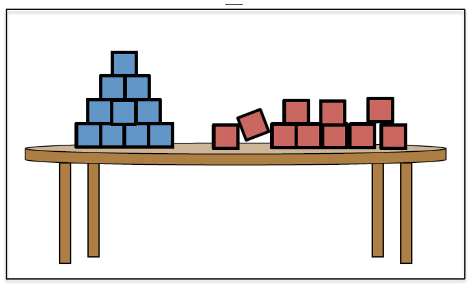
**Because it is ______________________________**

9



What will happen when the table shakes?



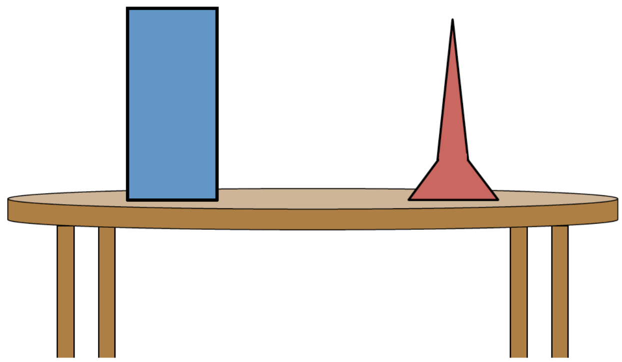


10

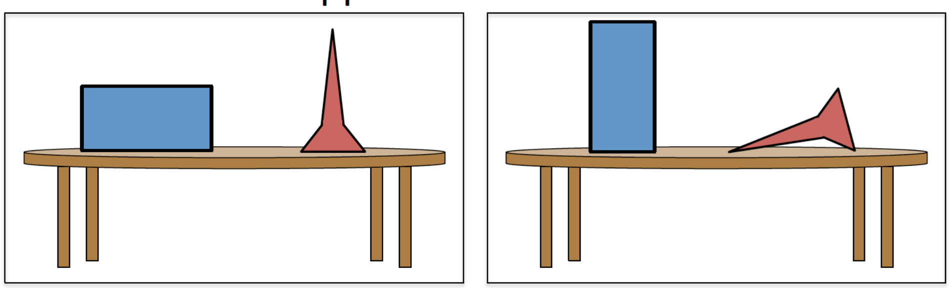
# Why did this one stay up?

**Because it is ______________________________**

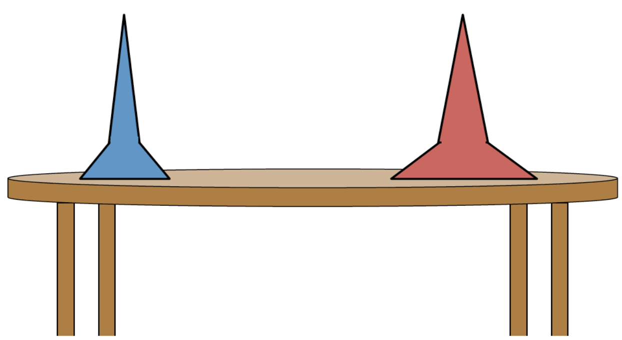
11



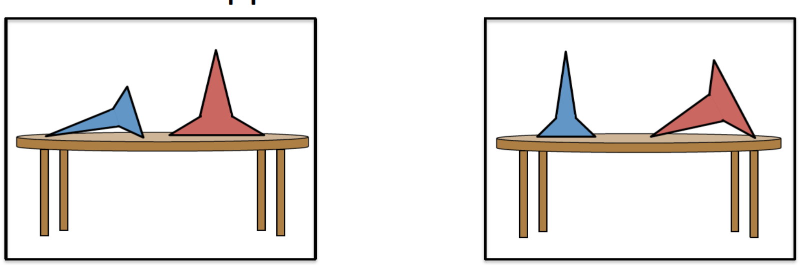
What will happen when the table shakes?

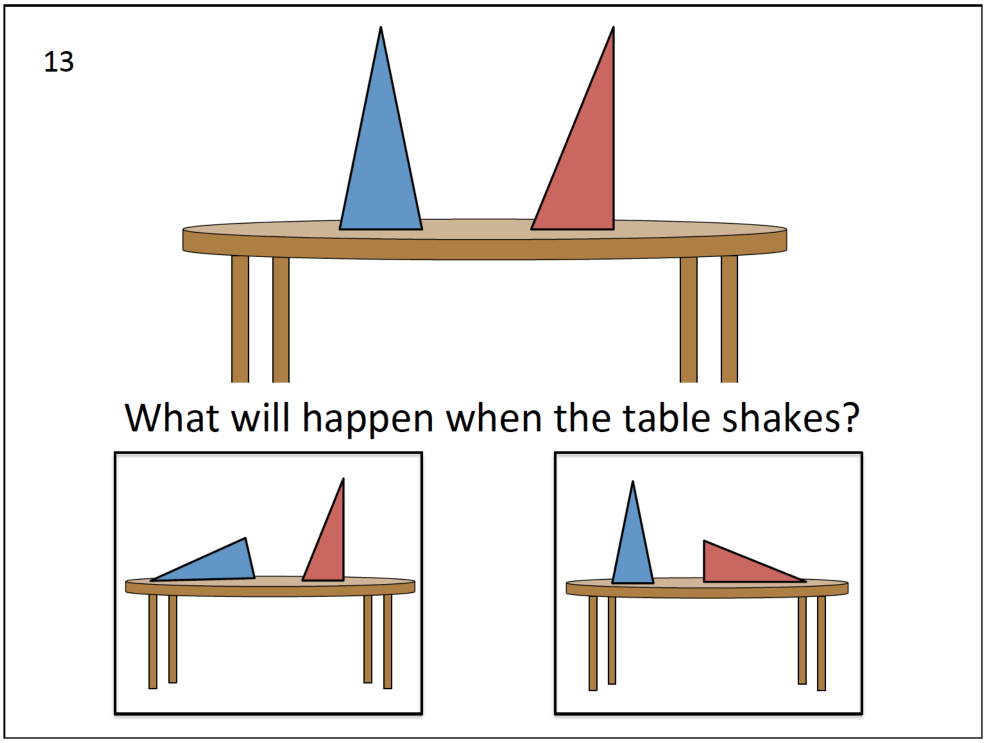


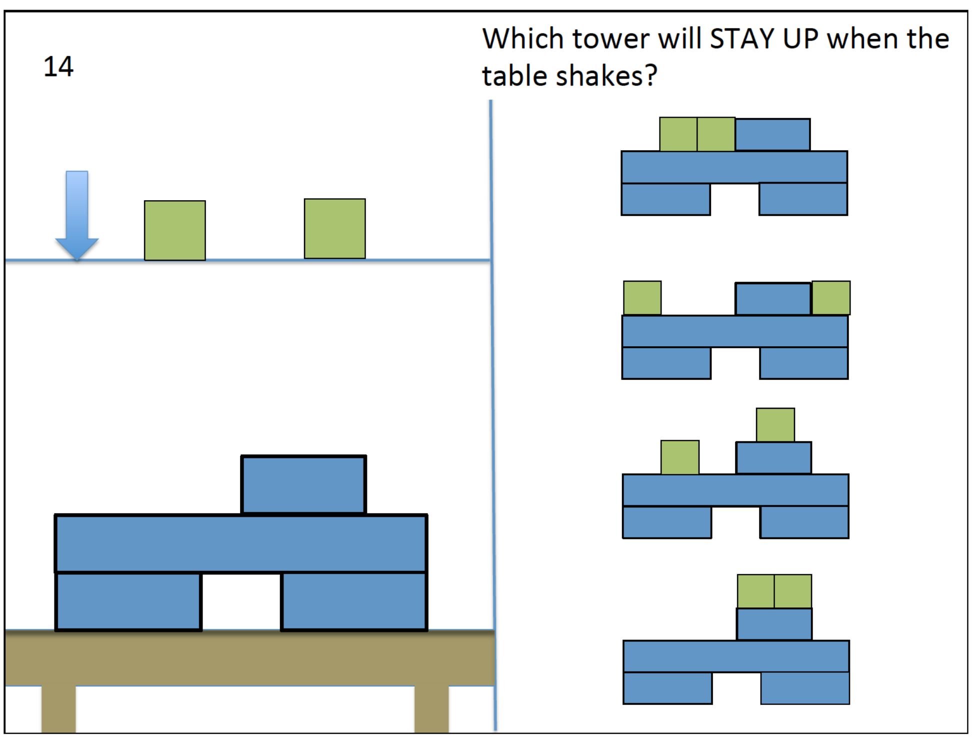
12



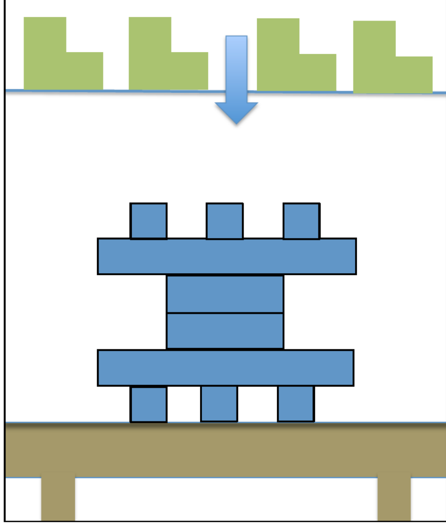
What will happen when the table shakes?



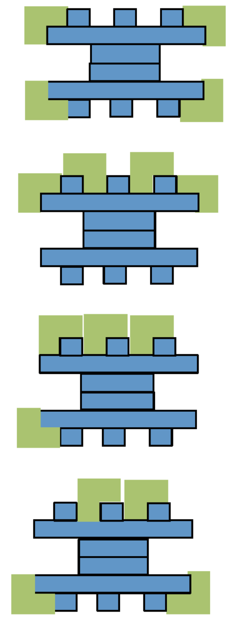

13
What will happen when the table shakes?


14
Which tower will STAY UP when the table shakes?

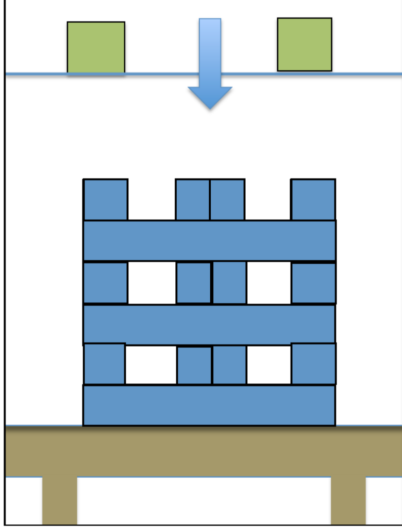
15



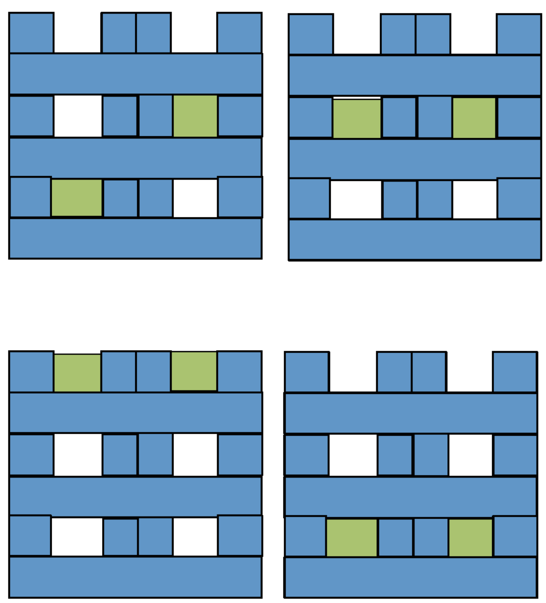
Which tower will STAY UP when the table shakes?



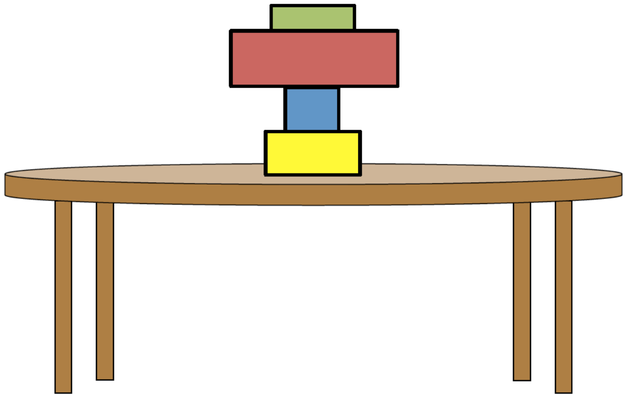
16



Which tower will STAY UP when the table shakes?



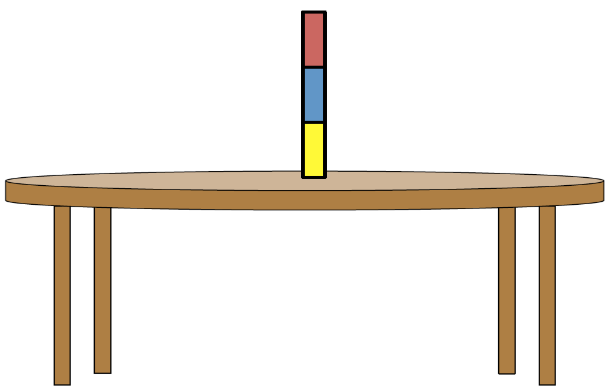
17



This tower will quickly fall when the table shakes. How could you make it STAY UP? Choose one.

a. Switch the red and blue blocks
b. Switch the yellow and blue blocks
c. Move the yellow block to the top

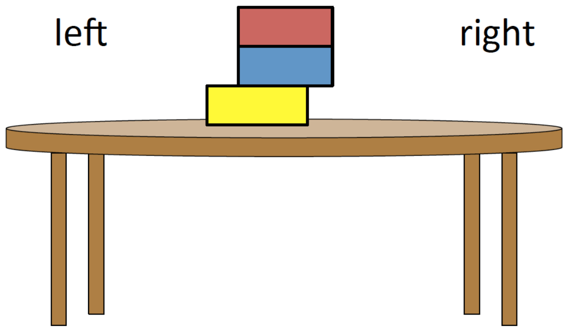
18



This tower will quickly fall when the table shakes. How could you make it STAY UP? Choose one.

a. Use a wider red block
b. Use a skinnier yellow block
c. Use a wider yellow block

19




This tower will quickly fall when the table shakes. How could you make it STAY UP? Choose one.

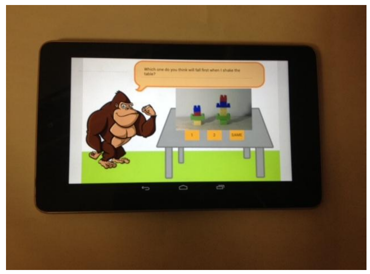
a. Move the yellow block to the right
b. Take the red block away
c. Move the red block to the left

# Appendix III: Enjoyment Survey for Experiment 2

## Tablet Condition

Name: ----------------------------------

How much did you like the game? (Circle one face)



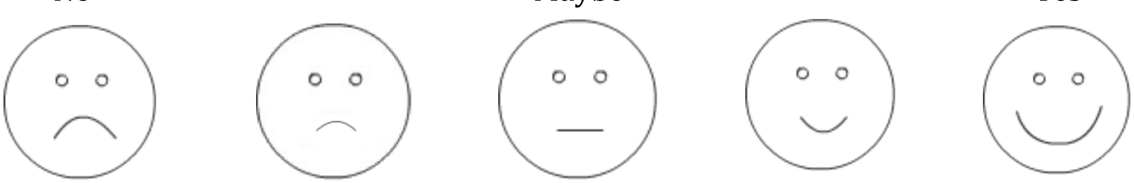
| I didn't like it at all | I didn't like it | It was OK | I liked it | I liked it very much |
|---|---|---|---|---|

Would you like to play it again?

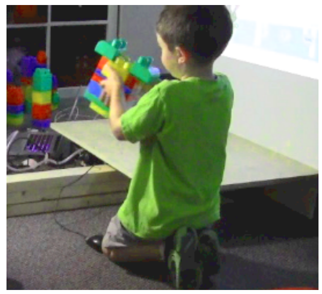
| No | | Maybe | | Yes |
|---|---|---|---|---|

Would you recommend it to a friend?

| No | | Maybe | | Yes |
|---|---|---|---|---|



How much did you like building your tower and testing it on the earthquake table? (Circle one face)



I didn't like it at all | I didn't like it | It was OK | I liked it | I liked it very much

# Appendix IV: Enjoyment Survey for Experiment 2

## Virtual Mouse Condition

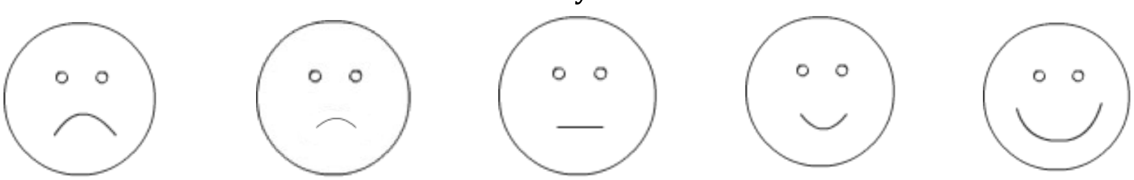
Name: ----------------------------------

How much did you like the game? (Circle one face)

I didn't like it at all | I didn't like it | It was OK | I liked it | I liked it very much

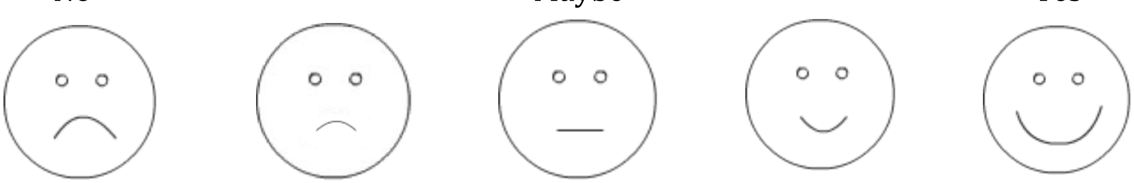
Would you like to play it again?

No | Maybe | Yes

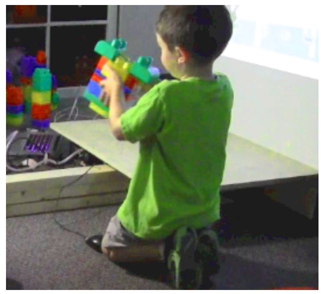


Would you recommend it to a friend?

No | Maybe | Yes



How much did you like building your tower and testing it on the earthquake table? (Circle one face)



| I didn't like it at all | I didn't like it | It was OK | I liked it | I liked it very much |
| --- | --- | --- | --- | --- |

# Appendix V: Enjoyment Survey for Experiment 3

Name: ----------------------------------

How much did you like the game? (Circle one face)

I didn't like it at all | I didn't like it | It was OK | I liked it | I liked it very much

Would you like to play it again?

No | Maybe | Yes

Would you recommend it to a friend?

No | Maybe | Yes

How much did you like building your tower and testing it on the earthquake table? (Circle one face)

I didn't like it at all | I didn't like it | It was OK | I liked it | I liked it very much