This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TLT.2017.2750673, IEEE Transactions on Learning Technologies

Hybrid Augmented Reality for Participatory Learning: The Hidden Efficacy of the Multi-User Game-based Simulation

Seungjae Oh, Hyo-Jeong So, and Matthew Gaydos

Abstract—The goal for this research is to articulate and test a new hybrid Augmented Reality (AR) environment for conceptual understanding. From the theoretical lens of embodied interaction, we have designed a multi-user participatory simulation called *ARfract* where visitors in a science museum can learn about complex scientific concepts on the refraction of light through full-body immersion using optical see-through AR glasses, projection-based AR, and gesture technology. In particular, we developed two different types of simulations for *ARfract*, namely a *game-based simulation* and a *non-game simulation* to explore how the order of different AR simulations influences the perceived usability, learning experiences, and learning outcomes. For the experiment, 10 dyads were randomly assigned to one of the two experimental conditions: 1) the game-to-non-game condition and 2) the non-game-to-game condition. The results indicate that the learners who experienced the game-based simulation before the non-game simulation performed better than did the other group with the reversed experience order. This paper also reports the usability and learning experience issues regarding the affordances of hybrid AR technologies. The major contribution of this proof-of concept research is that it articulates our understanding of how particular configurations (i.e. order) of the emerging technologies (i.e. hybrid Augmented Reality systems) and its use can lead to different learning outcomes.

Index Terms— Augmented Reality, wearable technology, embodied interaction, informal learning, refraction of light

1 INTRODUCTION

FOR the past decade, there has been a growing awareness of the important role of informal learning spaces, such as museums and science centers, and their role in extending learning experiences beyond formal school contexts [1], [2]. The nature of learning experiences in such informal learning spaces tends to be self-directed, emergent, relatively unstructured, and social. For instance, prior research about interaction patterns among visitors at science museums indicates that visitors tend to engage with exhibits for a limited time due to the lack of guidance and the myriad of choices of stimuli and information to attend to [3]. The public nature of exhibits and the diversity of vistors' profiles are other challenging issues for promoting interaction across exhibits and visitors.

Recognizing both the pros and cons of free-choice, emergent learning in a science center, we have been exploring the possibility of designing an interactive exhibit where multiple co-located visitors can interact through a participatory simulation. From the theoretical lens of embodied interaction, which foregrounds the interwoven nature of body, mind, and knowledge, we have designed a participatory simulation called *ARfract* with which visitors can learn about complex scientific concepts on the refraction of light (Fig. 1). The design aims of *ARfract* were 1) to encourage in-depth observations and discourse among multiple types of visitors such as doers, watchers and passers-by in a public learning space, 2) to accept natural body movement as an important input for deeper understanding from the perspective of embodied interaction, and 3) to naturally guide learners to develop sophisticated understanding about scientific phenomena.

This paper discusses the design, development and evaluation of *ARfract* regarding its efficacy on learning about abstract scientific concepts through bodily gestures and AR content. *ARfract* creates full-body immersion through hybrid augmented Reality (AR) visualization that leverages the affordances of optical see-through AR glasses, AR projection and gesture technology. We designed two different types of AR scenarios to deliver the refraction of light content, namely a game-based simulation and non-game simulation, and to explore how the different sequencing and level of exploratory learning through AR simulation can lead to better conceptual learning of complex science phenomena.

- *Seung Jae Oh is a doctorate student in the department of Creative IT Engineering, Pohang University of Science & Technology (POSTECH). 77 Cheongam-Ro. Nam-Gu. Pohang. Gyeongbuk. Korea 37673. E-mail: seungjae.oh7@gmail.com*
- *Hyo-Jeong So is with the Department of Educational Technology, Ewha Womans University, 52 Ewhayeodae-gil, Daehyeon-dong, Seodaemun-gu. E-mail: hyojeongso@ewha.ac.kr.*
- *Matthew Gaydos is with the Department of Technology and Society, SUNY Korea, Incheon, Yeonsu-gu, Songdo 1(il)-dong, E-mail: matthew.gaydos@gmail.com.*

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TLT.2017.2750673, IEEE Transactions on Learning Technologies

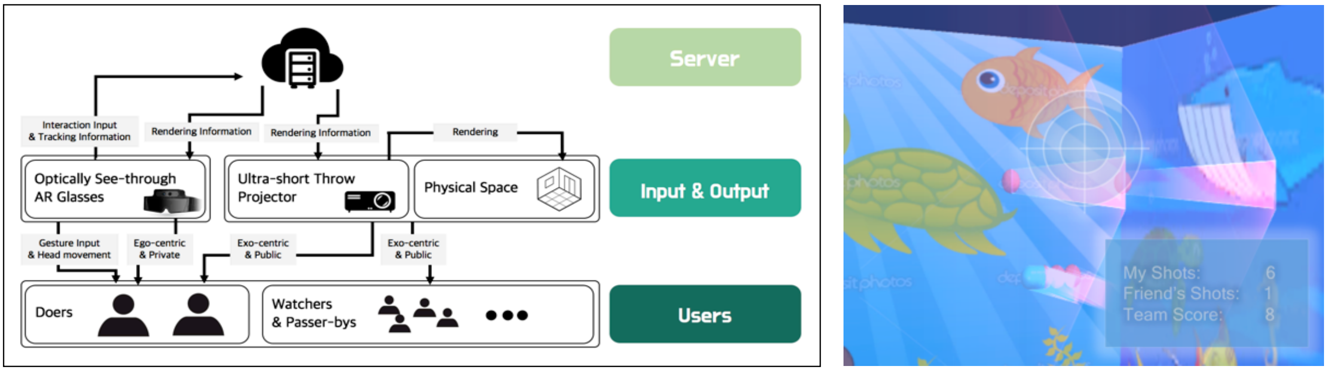

Fig. 1. (a) The infrastructure of the *ARfract* system & Visualization of the data flow (Left). (b) The screen shot of ego-centric view at AR glasses. Two pink balls & a blue ball with light trail represent the propagation of the lights figure in the caption (Right).

Central to our investigation was a question of how to use the AR system effectively, specifically regarding how to best order the sequence of the goal-based (game) and unguided (simulation) activities. While simulations and games can be effective tools for science learning [4], their effectiveness depends on appropriate design and use [5]. Differently structured participants experiences, even with the same tools, can lead to significantly different learning outcomes [6], [7].

2 Literature Review

2.1 Affordances of Augmented Reality and Learning

Milgram and Kishino [8] suggested a reality-virtuality continuum that encompasses augmented reality (AR), augmented virtuality (AV) and mixed reality (MR). The continuum is a continuous scale ranging from the completely real environment to the completely virtual environment. On this continuum, AR is positioned close to the real environment where users perceive real objects or phenomena with an additional layer of virtuality. Another widely-used definition of AR is from Azume [9] where AR is defined with three core characteristics: a) combining real and virtual, b) interactive in real time, and c) registered in 3-dimension (3-D).

As these classic definitions suggest, the core feature of AR lies in its visual representational power to combine real and virtual worlds in a 3-D space. In addition, a centricity continuum concerning users' viewpoints from egocentric to exocentric [10] is a useful framework to differentiate visual representational power between desktop-based simulation and head-mounted display (HMD) AR simulation. Here, egocentric refers to a viewpoint within the user (first-person view) whereas exocentric refers to a viewpoint from outside the user (third-person view). Desktop computers, as a two-dimensional medium on monitor displays, have an innate problem in presenting abstract scientific concepts, especially those that are three-dimensional in nature [11], [12]. Furthermore, since the control space of desktop computers is mostly world-referenced and two-dimensional, a desktop-based interactive simulation can only provide a two-dimensionally projected top-down perspective (exocentric) where symbolic manipulation (e.g. keyboard buttons or icons) may not be directly mapped to scientific concepts that require three-dimensional perceptions.

As an alternative, augmented reality extends the physical space with virtual information [13]. Hybrid AR that integrates wearable AR glasses and projected AR has the potential to provide tightly coupled exocentric (projected AR) and egocentric (AR glasses) viewpoints of physical space [10].

AR has enormous potential for learning, and various computing applications embracing AR have been widely applied in diverse learning contexts. Meta-reviews have already begun to synthesize the impact of AR for teaching and learning. For instance, Santos and colleagues [14] conducted a comprehensive meta-analysis of 87 research articles concerning AR learning experiences. They found a moderate effect size (d=0.56) in the performance of students who used AR-based approaches versus non-AR based approaches. They also identified three inherent advances of AR: a) real world annotation, b) contextual visualization, and c) vision-haptic visualization. In another meta review by Bacca and colleagues [15], they analyzed 32 articles (published between 2003 and 2013) on the use of AR in education settings, and found that AR was dominantly applied in higher education settings with relatively fewer studies in K-12 contexts. Further, they found that marker-based AR (e.g., using maker labels containing a colored or black/white pattern information registered by the AR application) was the most widely used technology in the previous studies, suggesting the need to explore other innovative AR technologies in future studies.

Some technical challenges in AR applications have been identified in the previous studies. For instance, Chang et al. [16] developed a mobile AR application to be used at an art museum. Users in their study, however, were always holding a mobile device to augment digital contents over the space, which makes it rather challenging for AR to be merged with flexible bodily gestures. Lindgren and Moshell [17] designed a mixed reality simulation environment projected on the floor to deliver planetary astronomy. In their study, while physical interactions with the digitally augmented environment provide meaningful and playful learning

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TLT.2017.2750673, IEEE Transactions on Learning Technologies

experiences, more complex scientific concepts that involve 3-D understanding are limited by the restricted visualization capability of a 2-D surface, inherent to projected AR.

Recently, hands-free, wearable AR technologies such as Google Glass have been applied in the education field. Research on such technologies is still scarce, however. From the thematic review of literature, Bower and Sturman [18] identified the unique affordances of wearable technologies in education under three themes: a) pedagogical uses, b) educational quality, and c) logistical and other implications. The affordances for pedagogical uses include in situ contextual information, recording, simulation, communication, first-person view, in-situ guidance, feedback, distribution, and gamification. Under educational quality, engagement, efficiency, and presence emerged as the key affordances of wearable technologies. Related to logistical issues, the potential of wearable technologies is related to freeing up space and hands-free access. As concrete cases of wearable AR for teaching and learning, Wu, Dameff and Fully [19] report a study of medical training where students participated in the simulation-based training scenarios that use Google Glass to provide a first-person view and recording function. Similarly, Coffman and Klinger [20] conducted a trial study in an educational psychology course where students were provided with Google Glass and used it to perform multiple activities such as taking pictures, video recording, and Internet access.

2.2 Effect of Sequencing in Open & Public Learning Experience

AR can be used as a tool to improve formal learning and to actively engage students [21]. As a new technology, however, research is still determining how to best take advantage of the learning affordances of the medium. AR can be challenging both in terms of its design and use when adapting it for education [17], [22]. As augmented reality technologies continue to develop and make their way into formal and informal learning settings, a primary agenda for research will be developing theoretical understanding and practical recommendations for the design of curricula that effectively uses the technology.

In informal learning settings like museums, the design of exhibits has been explored with regards to the narrative that museum attendees construct [23]. Creating exhibits that convey narratives of scientific phenomena is challenging, particularly compared to creating exhibits that convey historical narratives [3]. Interactive visualizations provide a possible means for helping people make sense of scientific phenomena, though issues remain with regards to ideal implementation, as guided or structured interaction must be balanced with participants' freedom [24].

How to structure learning experiences in informal learning settings is a complex issue. DeWitt and Storksdieck [25] conducted a review of research on museum studies and concluded that "field trips should provide a moderate amount of structure while still allowing for free exploration" (p.186). Gutwill and Allen [26] conducted a research study that aimed to find a sweet spot between choice and guidance in the Inquiry Games program designed to teach students about inquiry skills. By contrasting the efficacy of four different configurations of exhibits that differ in structure and collaboration among visitors, they found that structured and collaborative experiences led to the highest learning gains compared to the conditions that are spontaneous and individualized.

The study by Yoon and colleagues [27] aptly points out the tension between structuring learning to increase scientific understanding and preserving visitors' informal learning behaviors. In their study conducted with over 300 students at the science museum, the impacts of different combinations of scaffolds ranging from the AR digital augmentation to knowledge building scaffolds were investigated. Their study found a reverse relationship between learning scaffolds and informal behaviors (i.e., self-directed exploration and question-generation), implying that providing more structures tend to influence less informal learning behaviors, except when students collaborated in a small group.

Considering the complexity of optimizing informal learning experiences, this study aims to contribute to the existing body of literature about informal learning and sequencing/structuring learning activities. In particular, this study differentiates between game and simulation and explores a finer distinction of exhibit structuring, looking at the sequence or order in which a museum attendee should experience such structures. Though the order of instructional activities has been shown to influence learning in formal classroom environments, such sequencing has not been investigated in ill-structured, informal science learning settings (i.e., museums) [28], [29].

3 ARFRACT: AR SIMULATION ENVIRONMENT

In this section, we present the detailed description of the AR simulation environment called *ARfract* specifically developed for the current study. The *ARfract* is a room-sized simulation environment designed for a science museum context, leveraging the affordances of hybrid AR technologies. In *ARfract*, visitors can learn about abstract scientific concepts that cover the propagation of incident light and refracted light at a boundary of physical materials. We chose this topic mainly for three reasons. First, the phenomena related to the travel of light are mostly invisible to human eyes, making it difficult to understand [30]. While pencil drawings or 2-D simulation is a typical visualization approach used to teach this topic in school contexts, students still have a difficulty understanding this invisible complex phenomenon. *ARfract* attempts to alleviate this problem by representing how light travels in the virtual world space. Second, understanding the refraction of light involves changing mediums, which can be challenging to do in a real-world physical setting but which again, can be simulated in the virtual environment. Third, some visible light sources that

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TLT.2017.2750673, IEEE Transactions on Learning Technologies

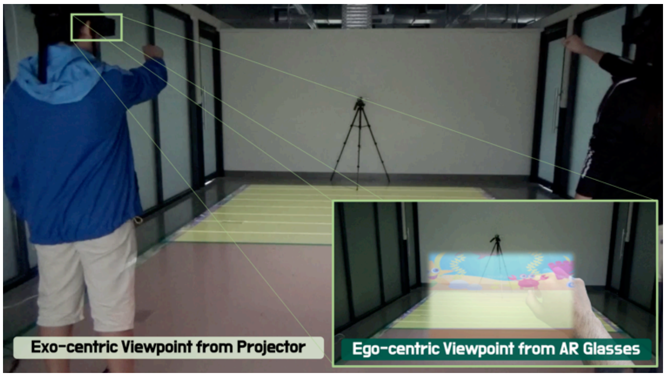

Fig. 2. Scene in a green box shows the first-person perspective. Tightly coupled with physical space, projected information and AR glasses information (hybrid AR) are presented simultaneously to the user.

Table 1
Design Consideration of AR Simulation Content Types

Design Considerations	Game-based Simulation	Non-game Simulation
Learning mode	Goal-oriented	Task-oriented
Expected level of collaboration	High	Low
Expected level of vicarious learning	Low	Moderate
Light trails on the floor as visual scaffold	No	Yes
Walking along the light trails	No	Yes
Interaction strategy	Backward planning	Forward planning

are useful for studies of refraction can be dangerous without supervision (e.g., lasers). An AR system can be a safer environment to convey such dangerous content in real-life settings.

The system architecture of *ARfract* includes the integration of three functional components: a) optical see-through AR glasses, b) projected AR on the floor, and c) physical space. As shown in Fig. 1 (a), the system I/O control information is transmitted to the system server, and the data are processed in real-time, distributed by the server to the visualization components.

3.1 AR Hardware Set-up

We used optical see-through AR glasses (Meta One) with 960 x 540 resolution and a 23-degree field of view. The AR glasses provide 9-axis inertial measurement with 360-degree head tracking and a 3-D time-of-flight depth camera to receive hand gesture input. A projector (Panasonic PT-CW331rea) was set up to present a top-down projection with a resolution of 1280 x 800 on an area of 3.2 m x 5.44 m. We used Meta One SDK and Unity 3d for application development and rendering, and Photon Unity Networking service to connect AR visualization components.

Specifically, the AR glasses played a significant role providing 1) personalized augmented information in an ego-centric viewpoint as presented in Fig. 1(b) (e.g., instructions, user interfaces, and simulation contents in the first-person perspective) to users, and 2) a receiver for gestural inputs. AR glasses allow both hands to be free from holding a visual display, which supports flexible bodily gestures. In addition, the transparency of the glasses-type AR affords the novel experience of a tightly coupled digital-physical world.

For the projected AR set-up, we used an ultra-short throw projector for simple installation and less occlusions, which is advantageous compared with a projector installation at the ceiling. The projected visualization provides users with the third-person perspective of the simulation among visitors around the installation, which offers a global awareness that helps the visitors better comprehend the simulated phenomena (Fig. 2) [10]. With the combination of both the project AR and wearable AR devices, the users are exposed to the multiple dynamic 3-D visualizations from both globally presented graphical data on the projected floor and locally projected transparent graphics from the transparent AR glasses.

The *ARfract* affords digital augmentation with both a first-person view (egocentric) and top down view (exocentric), which are controlled by a direct input, head tracker (ego-referenced control) shown in Fig. 3. This configuration (like the ego-referenced handle control of driving experiences with both car navigation and driving scene) is expected to provide users with both immersive and expressive learning experiences, which is advantageous compared with the desktop-based 2D simulation, which uses a top-down view (exocentric) and an indirect mouse input (world-referenced control) [10].

Through the integration of the optical see-through AR glasses and the projected floor, it is possible to combine individually augmented space and shared public space, thereby extending learning experiences toward a new hybrid AR realm (Fig. 2). In summary, leveraging the learning affordances of the AR technologies, the *ARfract* simulation is expected to a) accommodate multimodal learning, b) to provide opportunities for more interaction possibilities with visitors' free body and hand movements (e.g., embodied interaction with a digitally augmented space and social interaction among multiple co-located visitors), and c) to provide intuitive visual representations of abstract scientific concepts.

3.2 AR Simulation Content

For the experiment, we developed two different types of simulations for *ARfract*, namely 1) a game-based simulation and 2) a non-game simulation. Both the game and non-game simulation were designed to teach the refraction of light and related concepts. Also, they were both intended for use in an informal learning context, used the hybrid AR technology, and aimed to promote individual and social interactions using the technology as a mediator.

The design of *ARfract* includes two different types of social interactions to enhance communication and collaboration among users: 1) interaction between current users and 2) interaction among current users and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TLT.2017.2750673, IEEE Transactions on Learning Technologies

Fig. 3. First- person view (ego-centric, left) and top-down view (exocentric, right) in the *ARfract* simulation

bystanders. Interaction results of the current users are displayed on the floor as public information, and provide shared information among visitors so that they may socially interact with each other.

Also, there are several features for supporting embodied interaction. Participants using the *ARfract* wear optical see-through AR glasses and have opportunities to perform several metaphorical and intuitive bodily actions. For instance, participants rotate their body or head and perform a grab-and-release hand gesture to indicate that the light should travel in the direction they are looking at. The hand gesture indicates grabbing the light and releasing it to travel (metaphorical) toward the direction where the participants are looking at (intuitive). In this way, participants use their body as a reference to improve their simulated learning experiences. The detailed procedures how the simulation and game are used in the actual experiment are described in subsequent sections (Section 3.2.1 & 3.2.2). Below, we mainly focus on describing the core design features of each AR simulation type, which is also summarized in Table 1 for comparison.

3.2.1 Game-based Simulation

The game-based simulation consists of five stages with different materials or shooting locations. Basically, participants are asked to hit a target opposite to themselves with a beam of light. In Fig. 4(a), There are five different stages with different medium compositions through which the light passes: air-air (Stage 1), air-water (Stage 2), air-diamonds (Stage 3), water-air (Stage 4) and variations in the positions of each medium (e.g., water-air but closer to the boundary to increase the total refraction) (Stage 5). Difficulty levels of stages were designed to increase in the game-based simulation, and the analysis of interaction logs (Section 5.3.3) was conducted to examine this aspect of the design.

Collaboration is encouraged in the game-based simulation with two users working together as a team to shoot the target given limited chances. Here, both users wear AR glasses and can control the direction of light individually. The team score is presented to both users on their AR glasses. Distinct from the non-game simulation condition, both participants are able to control who shots the light and the light's direction, and thus are encouraged to predict the light's trajectories at different stages with various medium compositions. A stage ends when both participants successfully hit a white ball located beyond the boundary. The color of the ball change from white to purple when both participants successfully hit the ball. In this game, there are up to 15 chances as a group to hit their targets (see Fig. 1(b)). AR glasses displays information about the shooting counts of participants and total scores at each stage. All users experience the five stages in the same order.

The game-based simulation provides less scaffolding than the non-game simulation, letting users find their own solutions in exploratory and collaborative manners. That is, unlike the non-game simulation, the game-based simulation does not display the trajectory of recent light travel (Fig. 4(a)). Learning takes place implicitly while players make their own strategies to achieve the desired path of light travel (backward planning). The game-based simulation offers a series of visual scaffolds (i.e., light trails on the floor) that disappear right after an execution. Reflection and elaboration are encouraged as participants revise their previous execution and attempts to hit the target.

3.2.2 Non-game Simulation

The non-game simulation is characterized by the simple conversion of the 2D virtual simulation from the desktop environment (using WIMP) to the physical environment (using AR technologies). In the simulation without game elements, there are four designated locations (close or far from the boundary and two different materials) and two roles (shooter and observer) (see Fig. 4(b)). Interactions at different distances or with different materials serve as a method to enhance inductive learning [31], [32].

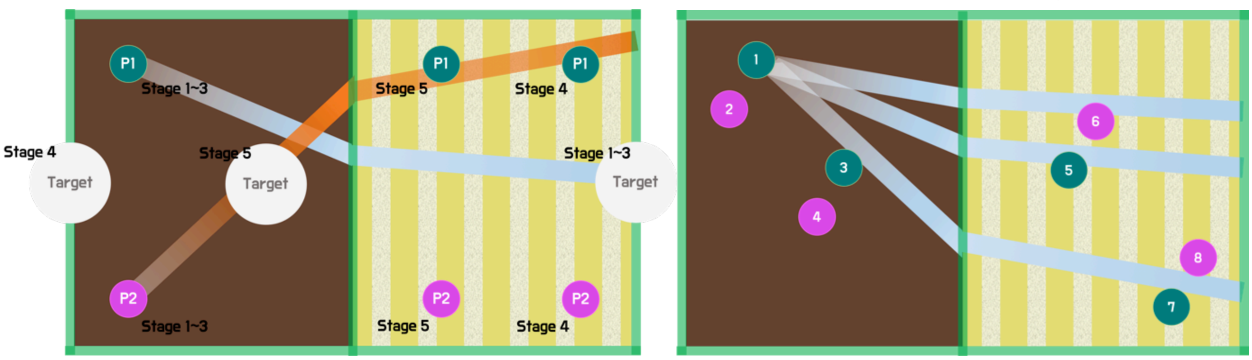

Fig. 4. (a) The stages of game-based simulation. Only the current paths of both users are presented. From Stage 1 to 3, a material of the other side is changed and a shooting location is changed at Stage 4 & 5 (Left). (b) Non-game Simulation. Green circles and red circles denote shooting position and observing position respectively (not visible during the simulation). Numbers in the circles represent an order of position changes. Three recent paths of a user at "shooting position #1" are presented on the floor (Right).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TLT.2017.2750673, IEEE Transactions on Learning Technologies

Users with AR glasses are instructed to stand on a shooting position (shooter) and an observation position (observer), and change their positions when each turn is over. There are 4 shooting and 4 observation positions for each participant (Fig. 4(b)). During each turn, users are instructed to try firing light in different directions. A user at the shooting position performs the grab and release gesture to fire light while an observer standing next to the shooter observe the interaction. A ball indicating the forefront of light and trajectory of light is presented to the users through the AR glasses (see Fig. 1 (b)). The light refracts when crossing the boundary between air and water. The simulation ends once users complete their eighth turn.

Before each turn, users are given the chance to walk along their recent light trails. The water side of the floor is lined with Velcro to slow down users' walking speed and to create a physical relation with the experience of traveling through a different medium. While walking on the floor, users experience several embodiment features. As they follow the light trail, they can rotate their body at the boundary of different materials, noticing the relationship between the incidental and refraction angles. Velcro on the floor disturbs walking movement speed, implying and reinforcing the phenomenon of why the light refracts at the boundary. In this regard, we intended that walking through the space, especially with tactile feedback, could have a significant influence on participants' embodied learning [12].

Compared to the game-based simulation, the non-game simulation offers more support (e.g., three cumulative light trails, rich bodily movement, and turn-taking). Users have more freedom to do the experiment based on the instantaneous feedback and feedforward mechanisms enabled by the simulation. The opportunities to observe and interact with a companion are designed to foster vicarious learning to happen between users.

3.3 Physical Space Design

As a last key component of the *ARfract*, the physical space is a rectangular-shaped room with the green-colored lines on the floor that provides implicit information to visitors such as: 1) a dividing line at the center to represent the boundary between the material with large refractive index and low refractive index, and 2) rectangular lines covering the projected area to discriminate the realm of virtuality and reality. Passers-by can easily notice where to stand and observe the interaction. We attached Velcro lines on half of the rectangular floor to represent materials with high refraction index where the speed of light slows down. As explained in the non-game simulation, users can walk along the Velcro-attached floor where walking speed is reduced, enhancing the metaphorical mapping between the action and the scientific phenomenon [33], [34]. To seamlessly integrate the physical space and virtual space, we aligned the camera matrix parameters of the AR glasses with the projected virtual space provided by Unity 3D.

4 METHOD

4.1 Research Rationale and Questions

The goal for this research is to articulate and test a new hybrid AR environment for conceptual understanding in science. The main purpose of this paper is to provide both rich descriptions of the AR environment's configuration and its impact on learning experiences and outcomes. Our design, that is an AR interactive simulation that presents scientific phenomena about the travel of light and its refraction at the boundary between two media, is only one possible application of the AR technologies in education. Although AR, games, and simulations have all been shown to promote science conceptual learning in different contexts, combining them is non-trivial, and this project tests the order of two configurations (game/non-game). Theories of embodiment as well as game-based learning may be generally useful for informing elements of a simulation, but such theories must be appropriately applied through design in order to take advantage of the technologies available. Thus, the fundamental goal of the research with respect to learning outcomes is to answer the following research questions:

Research Question 1: Do experiences with *ARfract* help participants learn the complex science phenomenon about the refraction of light?

Research Question 2: Does the presentation order of AR simulations (the game-to-non-game condition vs. the non-game-to-game condition) influence learning outcomes?

Research Question 3: Does the presentation order of AR simulations (the game-to-non-game condition vs. the non-game-to-game condition influence the perceived usability and learning experiences with *ARfract*?

By answering these questions, we expect to enhance our understanding of how particular configurations of AR installations result in different learning outcomes in informal leaning spaces using the emerging technologies (i.e., projected AR and AR glasses).

4.2 Experimental Design

4.2.1. Participants and Research Conditions

We recruited 10 mixed-gender dyads (20 participants, 14 males, 6 female) who ranged in age from 12-14 years old. A pre-survey was administered to identify the participants' backgrounds. The pre-survey data indicated that the participants' subjective ratings about their interest in science learning was M=3.75 (SD=1.18) and their learning motivation was M=4.25 (SD= 0.68) on a 5-point Likert scale. While only two participants had prior experiences with mixed reality applications, surprisingly, all participants reported that they had experienced at least one gesture-based simulation (e.g., Nintendo Wii, gesture-based games).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TLT.2017.2750673, IEEE Transactions on Learning Technologies

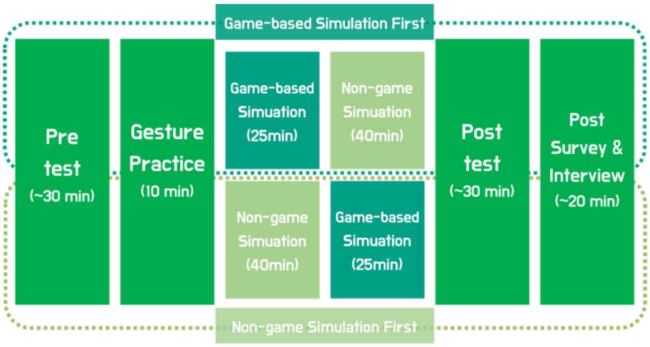

Fig. 5. Experiment procedures

We conducted the experiment in a lab setting. While *ARfract* was designed to be used in a science center context, it is important to note that we chose a lab setting for this particular experiment due to the number of variables and technical set-up (e.g., floor design, AR equipment set-up) that we wanted to control in a less intrusive setting. For the experiment, the 10 dyads were randomly assigned to one of the two experimental conditions according to the Experience Order factor (ExpOrder): 1) the game-to-non-game condition where the participants experienced game-based simulation first and then non-game simulation (Game First, GF), and 2) the non-game-to-game condition where the participants experienced the two types of simulation in a reverse order (Non-game First, NF). At the end of the experiment, all participants completed a post survey regarding their perceived usability and learning experiences. They also took a pre-test and post-test measuring their understanding of scientific concepts related to the refraction of light. Interaction logs were recorded in the database throughout the session. The types of interaction log data include the number of shooting, game scores, etc.

4.2.2 Procedures

Fig. 5 provides an overview of the experiment procedures: 1) to complete a pre-test with 15 questions about the refraction of light & total reflection and personal information (pre-test), 2) to practice firing gestures with guidance (practice), 3) to experience the AR simulation in a different order depending on their assigned experimental condition (experience), 4) to take a post-test measuring learning gains (post-test), and 5) to complete a post survey and semi-structured interview about overall learning experiences (post-survey & interview).

4.2.3 Gesture Practice

The gesture practice session occurred at the start of the experiment after the participants completed the pre-survey. The session lasted for approximately 10 minutes to help the participants feel comfortable with the core gestures in the *ARfract* system such as grabbing and releasing to initiate the travel of light. Participants could move their body or head to control shooting directions. The research facilitators carefully guided the participants until they could comfortably fire light in a desired direction.

4.2.4 Knowledge Test

The pre-test contained 9 multiple-choice and 6 short-answer questions that were designed to assess the participants' prior knowledge about light refraction and related conceptual understanding. The test items were co-designed by the research team and the science teacher. The test items contained various types of questions, ranging from knowledge and comprehension (e.g., how light refracts between various medium compositions) to application of complex scientific concepts (e.g., total reflection, Snell's law, and Fermat's principle).

To minimize learning effects between pre-test and post-test, the post-test design was a modified version of the pre-test. We modified the sequence of the test items and the item statements including questions about the different mediums and different directions of light travel. The content validity was confirmed by the physics education teacher. The post-test was used to evaluate the participants' learning performance after the experiment. In both tests, multiple choice items were scored as correct (1 point) or incorrect (0 point). Short answers were scored as correct (1 point), partially correct (.5 point) or incorrect (0 point). Two researchers cross-checked the test results for reliability.

4.2.5 Post-Survey: Usability and Learning Experiences

We adapted the questionnaire from [35] to measure perceived usability and learning experience. The questionnaire consists of 22 items in five factors: a) *overall perceptions of learning experiences* (5 items), b) *self-engagement* (5 items), c) *quality of embodied learning* (4 items), d) *usability of the AR simulation* (4 items), and e) *overall satisfaction* (4 items). Responses were measured on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

4.3 Data Analysis

R version 3.2.2 was employed to perform statistical analyses, and several R packages were used for visualizations. To discriminate statistical significance, a *p*-value less than .05 criteria and two-tailed tests were used. Descriptive statistics and paired-samples t-test were conducted to analyze the effect of *ARfract* on learning performance. Analysis of Covariance (ANCOVA) was also performed to identify the effect of the presentation order on the learning gains. Levene's test was used to examine whether the error variances were equally distributed between independent variables. For the post-survey data, an independent two-sample t-test was performed to compare the scores in the five factors between the two ExpOrder conditions. Normality assumptions were checked using the Shapiro-Wilk test, which confirmed that all data were sampled from the normally distributed population. Parametric tests, therefore, were used in the subsequent data analysis. Two dyads, however, were excluded from the data analysis because they lost concentration during the experiment and could not complete the post-test and post-survey. Interview data were transcribed and analyzed by the researchers based on the emerging themes through a

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TLT.2017.2750673, IEEE Transactions on Learning Technologies

TABLE 2
Descriptive Statistics for Paired Samples t-Test

	N	Mean	SD	SEM
Pre-test	16	6.88	2.34	0.59
Post-test	16	9.75	1.92	0.48
Difference		2.88		

constant comparison method. For the interaction log, an ANOVA was used to distinguish the variances from different game stages (Stage 1 to 5), and to compare the log of different simulation types (game-based simulation & non-game simulation). We did not conduct other analyses on the demographic variables (e.g., gender, age) since the sample size was small and participation was not balanced for these variables.

5 Results

Detailed results of the data analysis and statistical procedures are described in the following section according to each research question.

5.1 RQ1: Do experiences with ARfract help participants learn the complex science phenomenon about the refraction of light?

The overall analysis indicates that the participants had significant learning gains from the *ARfract* experiences (see Table 2). To test the overall effect of learning with *ARfract* on conceptual understanding, we aggregated data from both conditions (i.e., Non-game First (NF) and Game First (GF)). The mean value increased by 40 percent from 6.88 in the pre-test to 9.75 in the post-test. The maximum score of the test was 15 points. Among 16 students, only one student had a negative learning gain (Gain = -1) and another student had no learning gain. The remaining 14 students scored more than three points higher in the post-test compared their score in the pre-test. Paired-samples t-test with pre-test and post-test (n=16) showed a significant increase in the test scores ($t(15) = 5.97$, $p < 0.00$).

The *ARfract* experience in both Experience Order (ExpOrder) conditions turned out to be effective in learning about the light refraction. A paired-samples t-test was performed on the pre-test and post-test scores of each ExpOrder conditions (n=8 for each) to verify the effect of *ARfract* on learning outcomes (Table 3). First, a paired-samples t-test was conducted on the test scores of the GF condition where the dyads did the AR game-based simulation first and then moved to the non-game simulation. The analysis showed significant increase in the test scores from pre-test to post-test ($t(7) = 7.75$, $p < 0.00$). Similarly, a paired-samples t-test on the test scores of the NF group was performed. The analysis showed a statistically significant difference between pretest and post-test scores ($t(7) = 3.153$, $p = 0.01609$).

5.2 RQ2: Does the presentation order of AR simulations influence learning outcomes?

In RQ1, different learning gains were observed from different simulation experience orders. There was a need to compare the learning gains from the ExpOrder conditions, and to compare the changes between pre-test and post-test with the reduced error variance and the minimized bias. Hence, an ANCOVA was performed with the post-test score as the dependent variable, the pre-test score as a covariate, and the independent variable, the ExpOrder with two levels, as the fixed factor (Table 4). A correlation between the pre-test and post-test was high enough to apply ANCOVA (r =.608). Levene's test on the error variances of ExpOrder levels was not statistically significant ($F(1,14)$ =.037, p =.85). Several assumptions of ANCOVA were tested in a strict way, and it was concluded that the data set was suitable for ANCOVA.

The ANCOVA result suggested that there was a significant difference between NF and GF groups on the post-test scores ($F(1,13) = 10.54$, $p = .034$, $\eta^2 = 0.301$, indicating that the learning performance of GF participants (adjusted mean = 10.67) was significantly higher than that of NF participants (adjusted mean = 8.83). Thus, it was concluded that although students' final scores were similar, the condition (Game First or Non-game First) resulted in different learning gains, essentially favoring the game-first condition.

TABLE 3
Descriptive Statistics for Paired T-Test for ExpOrder

Group		N	Mean	SD	SEM
Game First	Pre-test	8	5.81	2.20	0.78
	Post-test	8	9.94	2.06	0.73
	Difference		4.12		
Non-Game First	Pre-test	8	7.94	2.08	0.74
	Post-test	8	9.56	1.90	0.67
	Difference		1.62		

TABLE 4
Descriptive Statistics for ANCOVA Test Result

Group	N	Mean	SD	Adjusted Mean	F
GF	8	9.94	2.06	10.67	10.54*
NF	8	9.56	1.90	8.83	

* $p < .05$

5.3 RQ3: Does the presentation order of AR simulations influence the perceived usability and learning experiences with ARfract?

5.3.1 Post-survey Data

Table 5 shows the descriptive statistics of the post-survey comparing the GF and NF groups in the five factors that were explained earlier in Section 4.2.5. Among the five factors, GF students tend to have high responses for Factor 1 (*Overall perception of learning experience*) and Factor 3 (*Quality of embodied learning experiences*) compared to NF students. Factor 1 was designed to evaluate whether the simulation experience with *ARfract* was not

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TLT.2017.2750673, IEEE Transactions on Learning Technologies

TABLE 5
Independent Samples T-test of Post-Survey

Post Survey Factors Items	Game First **Mean**	**SD**	Non-game First **Mean**	**SD**	**t**
Factor 1: Learning Experience (25)	**24.00**	**1.51**	**22.50**	**1.85**	**1.77′**
Easy to learn	4.63	.52	4.63	.52	.00
Fun to learn	4.88	.35	4.75	.46	.61
Effectiveness in learning	4.88	.35	4.25	.46	3.03*
Effectiveness in social learning	4.75	.46	4.38	.52	1.53
Fun in social interaction	4.88	.35	4.50	.53	1.66
Factor 2: Self-Engagement (25)	**20.38**	**3.07**	**18.63**	**1.85**	**1.38**
Immersive	4.13	.64	3.63	.74	1.44
Presence of myself	4.25	.71	3.38	.74	2.41*
Presence of others	3.38	1.19	2.88	1.13	.86
Distraction in attention (-)	4.13	.83	4.25	.46	-.37
Distraction in thinking process (-)	4.50	.53	4.50	.53	.00
Factor 3: Embodied Learning (20)	**18.75**	**1.39**	**17.25**	**1.28**	**2.25***
Motor memory	4.88	.35	4.63	.52	1.13
Effectiveness in understanding	4.75	.46	4.38	.52	1.53
Naturalness of bodily action	4.63	.52	3.63	1.06	2.40*
Action-thinking conflict (-)	4.50	.76	4.63	.52	-.39
Factor 4: Usability (20)	**17.00**	**2.39**	**15.25**	**3.37**	**1.20**
Easy to use	4.13	.83	3.63	1.06	1.05
Learnability	4.63	.74	4.13	.99	1.14
Controllability	4.25	.71	3.88	.99	.87
Comfortability of screen size (-)	4.00	1.07	3.63	1.19	.66
Factor 5: Satisfaction (20)	**19.38**	**1.41**	**19.25**	**1.04**	**0.20**
Enjoyment	4.88	.35	4.88	.35	.00
Recommend to friends	4.88	.35	4.63	.52	1.13
Intention to use	4.88	.35	4.88	.35	.00
Intention to participate	4.75	.46	4.88	.35	-.67
Total (110)	**99.50**	**8.57**	**92.88**	**6.83**	**1.71**

* $p< .05$ & ′ $p< .10$

only pleasant but also effective in learning with others. The mean score of Factor 1 showed higher ratings for the GF group (M=24.00) compared to the NF group (M=22.50), and the difference, tested by an independent two-sample t-test, was marginally significant ($t(14) = 1.17$, $p = .098$). Factor 3 represents the embodiment in the learning processes, which covers naturalness in bodily actions, action-thinking conflict, motor memory, and effectiveness in learning. As shown in Table 5, an independent two-sample t-test revealed significant differences in the mean values in Factor 3 between GF (M=18.75) and NF (M=17.25) groups ($t(14) = 2.25$, $p < .05$).

The other three factors include *self-engagement* (Factor 2), *usability of the AR simulation* (Factor 4), and *overall satisfaction* (Factor 5). These factors showed no statistical significance in the ExpOrder from an independent two-sample t-test, suggesting that the participants in both groups had experienced a similar level of engagement and usability when using the *ARfract*. Also, the mean values in Satisfaction were 19.38 for the GF group and 19.25 for the NF group (max = 20.00), indicating that the overall learning experience with *ARfract* was highly pleasant in both types of AR simulations.

5.3.2 Interview Data

The analysis of the interview data revealed four major themes about the perceived usability and learning experiences with *ARfract*: 1) pros and cons of wearable devices, 2) satisfaction with AR experiences, 3) self-directed and exploratory learning experiences, and 4) embodied interaction:

1) *Pros and cons of wearable devices*

Most participants wearing spectacles said, "The AR glasses are too bulky and heavy"; however, some reported that the AR sensor and head-tracking functions supported their learning experiences because they did not have to care about controlling where they looked. Due to the lack of sensor fidelity, the augmented graphics and real-world position were sometimes decoupled and misaligned.

2) *Satisfaction with AR experiences*

Participants felt that the invisible and intangible light seemed to exist and some even reported that, "It seems that I am really present in the virtual world." Most

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TLT.2017.2750673, IEEE Transactions on Learning Technologies

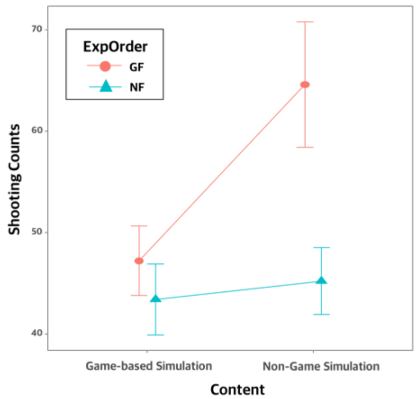

Fig. 6. A plot of *Shooting Counts* with respect to ExpOrder & Content

participants were satisfied with the design features regarding the display of light (i.e., the trajectory of light and its refraction through different mediums).

3) *Self-directed and exploratory learning experiences*
Many participants reported that the simulation was helpful in learning because they could manipulate the simulation with no restriction and at their own pace. The simulation afforded enough time and opportunities to reflect and explore. One participant described that, because the augmented contents were private, he felt more relieved and comfortable to interact. These findings are related to the private nature of head-mounted optical AR, which may be an important factor in the design of interactive simulations with AR glasses.

4) *Effectiveness of embodied experiences*
Overall, the gesture interaction where the participants were able to control the direction of light influenced their satisfaction, as they had the autonomy to manipulate phenomena and artifacts in ways they could not without the technology. Some participants reported moving their body while engaged in the simulation helped them to better comprehend the scientific concept. The participant commented that "Velcro at the water side disturbs walking, and it was well designed to convey the fact that speed of light slows down inside the water". Also, another participant said, "Turning my body to follow the light trail at the boundary of different materials helped me to realize differences between incidental angle and refraction angle". Our design of embodied interaction seemed to affect students' cognitive understanding through effectively mapping body movements and scientific phenomena, increasing the congruence between the two.

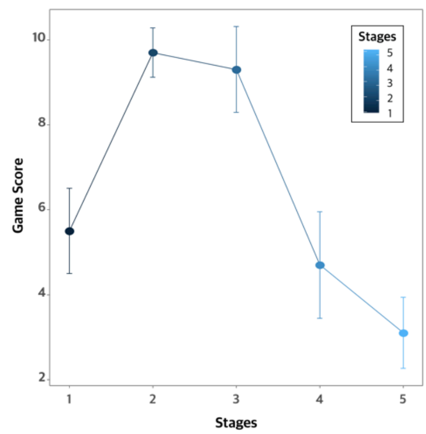

Fig. 7. A plot of *Game Score* with respect to Stages

5.3.3 Interaction Log Data

We used the interaction log data to investigate user behaviors according to the two simulation types (Content) and the presentation order of conditions (ExpOrder). The recorded interaction logs of both simulations were carefully inspected to discover meaningful findings about user behaviors and to confirm how the design considerations were realized in practice. Several measures like 1) *Shooting Counts* and 2) *Game Score* were calculated from the logs.

1) *Shooting Counts*
A two-factor mixed-design ANOVA was performed with *Shooting Counts* as a dependent measure and ExpOrder & Content as two independent variables. For the main effects, ExpOrder was marginally significant ($F(1,8) = 4.784$, $p = .06$) and Content was significant ($F(1,8) = 10.811$, $p = .011$). Furthermore, the results showed that there was an interaction effect between ExpOrder and Content, which was statistically significant ($F(1,8) = 7.137$, $p = .0283$). We performed simple-effect analysis on both independent variables (see Fig. 6). For the participants of GF condition, there was an increase in *Shooting Counts* in the non-game simulation condition (M = 64.60, SD = 13.85) compared to the game-based simulation (M = 47.20, SD = 7.66). In other words, GF participants (M = 64.60, SD = 13.85) made a significantly higher number of shoots than did NF participants (M = 45.20, SD = 7.40) in the non-game simulation condition (see Fig. 6). This implies that the structuring of game-based simulation prior to the non-game simulation promoted more active engagement and that engagement was sustained in the non-game simulation condition.

2) *Game Score*
Even though there was no significant difference between ExpOrder on *Game Score*, to verify whether the game-based simulation was consistent with our design intention, we performed a one-way repeated measure ANOVA with Stage as an independent variable and *Game Score* on the gamed-based simulation as a dependent variable. The result shows that *Game Score* was statistically different across the five stages in the game-based simulation ($F(4,36) = 8.373$, $p < .00$). From pair-wise comparisons between Stage, only the score at Stage 5 (M =3.1, SD = 2.6) was significantly lower than the scores at Stage 2 (M = 9.7, SD = 1.8) and Stage 3 (M = 9.3, SD = 3.2) (see Fig. 7).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TLT.2017.2750673, IEEE Transactions on Learning Technologies

6 Discussion

6.1 Hidden Efficacy of Game-based Simulation

The primary goal of this research was to develop and test an AR simulation-based learning environment that can be installed at a science museum. Recognizing the characteristics of informal learning environments where many visitors are constantly looking around the space, it is inevitable for them to experience the sequence of installations. To examine the effect of experience order on learning outcomes, we designed two types of simulation (i.e., game-based and non-game simulation) about the refraction of light. In the laboratory setting with 10 dyads, we attempted to examine whether the experience of two types of simulation in the *ARfract* environment could improve learning outcomes, perceived usability, and learning experiences.

Regarding the measures of learning outcomes from the different presentation order of AR simulations, the students who experienced the game-based simulation before the non-game simulation performed better than did the other group with the reversed experience order. This finding is rather surprising since the students from both groups reported similar levels of engagement, usability, and satisfaction about the AR simulation experience in the post-survey. For this result, we speculate that motivation toward explicit goals could be one potential reason that might have affected the different cognitive gains between the two groups. In addition, the group of students that experienced the game-based simulation first more actively engaged in the subsequent non-game simulation with 142.92 % of shooting counts compared to the reversed experience order. The higher engagement in the non-game simulation can be another potential evidence of motivation. Also, we expect the motivation resulted higher ratings on the *effectiveness in learning* item of Factor 1 (*Learning Experience*). The impact of motivation in game-based learning has been studied in broad fields. Further, prior research studies have suggested that motivation itself can serve as an intrinsic impetus for learning in open-ended informal learning environments [36], [37], [38]. One of the main differences between the game-based and non-game conditions was the existence of explicit goals to achieve. It is possible that the groups in the GF group were motivated to achieve the goals (hitting the targets within a given number of attempts) with the game elements (e.g., points, stage) in the simulation. Overall, this study suggests that having visitors experience exhibits with explicit goals and game elements may lead to better cognitive gains than having them experience exhibits without such elements.

Another important finding is that while cognitive gains were higher with the GF group than the NF group, both game-based simulation and non-game simulation are shown to be effective in enhancing scientific understanding as seen in the significant improvement between pretest and posttest. We attribute this positive outcome to the unique affordances of hybrid AR technologies employed in this study such as: a) the provision of both egocentric and exocentric views of complex science phenomena, b) the congruence between bodily guests and science concepts being learned, and c) the public and collaborative nature of the simulation. This speculation is supported by the high subjective ratings reported in the post-survey related to these items such as *effectiveness in learning, motor memory, and fun in social interaction*. This finding is also consistent with the previous research studies that examined the impact of AR technology on enhancing learning about abstract and spatial scientific knowledge [39], [40].

6.2 Hardware and Interface Design

This study also provides some implications about the design of AR simulation with wearable devices and gesture technology. The participants were satisfied and rated the overall learning experiences with *ARfract* highly. The learning experience with two simulation types encompassed a wide range of technical features (e.g., transparent AR glasses on the projected floor), multiple types of interactions (e.g., social and embodied interaction), and learning strategies (e.g., game-based learning and virtual experiment). However, some technical limitations of AR glasses were found in the survey. For instance, the score on *screen size comfortability* was the lowest, and we associate the potential reason for this with the nature of AR glasses as they cover only a small central region (23 degree) of the visual field. While we tried to compensate for this usability limitation by using the AR projection methods as Benko et al. [41] did, users still perceived the limited capacity of visualization in AR glasses. Although the transparency of AR glasses supported the participants' ability to see each other, the *presence of others* was reported as the lowest among Factor 4 (*Usability*). This supports our observation that there was a high variability in the social interaction across groups. While some groups in the experiment actively communicated verbally and through bodily gestures, some groups showed an extremely low level of such interactive patterns.

Embodiment is another important design feature of our AR simulation. Wearable AR glasses can effectively support embodiment because users do not have to hold a mobile display for augmentation. The wearable visualization in AR glasses allows users to freely look around the exhibit space because of the sensor-based head tracking feature. We chose single Inertial Measurement Unit (IMU) head tracking equipped in the AR glasses for development. Some participants, however, reported problems with the flow of the AR graphics, which we plan to improve with a tracking method that offers more accurate whole-body tracking using passive infra-red sensors or multiple IMUs. Improvements in tracking are likely to increase the usability of visualization allowing for more freedom in bodily actions and interaction possibilities.

Previously, embodied learning with AR technologies mostly relied on the projected AR techniques [12], which may impose some issues like unwanted disclosure or occlusion of information. In our design of *ARfract*, we

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TLT.2017.2750673, IEEE Transactions on Learning Technologies

used the optical see-through AR glasses to display private information and the AR projected floor to display public information. We speculate that such a seamless integration of public and private information might help users better engaged during the simulation, as seen in high scores in the post-survey such as the *intention to use* & *intention to participate again* items of Factor 5 (*Satisfaction*) and distraction in the *attention* item of Factor 2 (*Self-Engagement*). While the AR glasses used in this study interweave the multiple AR technologies, it was only possible to partially solve the occlusion problem with the AR glasses used in this study. The system was not wireless, making it still difficult for users to move and act at the same time.

Embodiment is another important design feature of our hybrid AR simulation. In the simulation, the participants were asked to walk along the light trails where the water side of the floor was lined with Velcro to slow down users' walking speed. The participants were also asked to turn their body at the boundary of two different materials to experience how the light refracts. The series of such metaphorical mapping in the *ARfract* simulation environment was designed to create the congruency between full-body movement and the phenomenon projected on the floor, which can consequently reduce *action-thinking conflict*. The data on Factor 3 (E*mbodied Learning*) in the post-survey suggests that embodied interaction in the *ARfract* had a positive influence on learning experiences, which is consistent with the existing literature that emphasized the power of embodied metaphorical mapping [33], [34], [42]. The rating on the naturalness of bodily action and some elements of usability, however, were rather low compared to other factors, which could be further improved with design features that naturally support full-body movement and visual information.

6.3 Implications on the Design of AR Simulation in Informal Learning

One of the main goals of this research was to prove the potential of hybrid AR simulation as an interactive exhibit in informal learning settings such as museums and science centers. Interactive exhibits are defined as "those in which visitors can conduct activities, gather evidence, select options, form conclusion, test skills, provide input, and actually alter a situation based on input" [43]. As implied in this definition, promoting inquiry-based thinking and skills through interaction with exhibits is an essential goal. Recently, with the advances of interactive technologies, exhibit design has incorporated the affordances of interactive technologies, as seen with the installation of AR-based and gesture-based exhibits in museums and science centers. However, our analysis of visitor interaction patterns in another study [44] revealed that visitors tend to show short engagement even with interactive exhibits.

Indeed, how to increase visitors' prolonged engagement with exhibits beyond superficial mindless interaction has been a long-lasting question in museum studies. Increasing technological interactivity in exhibits can be one way to engage visitors. Allen [3], however, suggests an inherent dilemma in interactive exhibits, as multiple interactive features overwhelm visitors and lead to disruption in the phenomenon being displayed. That means both too few or too many interactive features in exhibits may fail to engage visitors.

Similarly, another challenge in exhibit design is to decide on how much choice and structure should be given to visitors. In this study, we embedded the varying degrees of structure and choice in the two different types of AR simulation, with different design elements such as visual scaffolds, intended social interaction, and gamified activities. Our overall findings suggest that while both simulation types were effective for promoting cognitive gains, the game-based simulation appears to have a better optimal level of structure by supporting the participants to be engaged with the goal-based learning (game), yet allowing freedom to collaborate, experiment, and reflect. Such inquiry-related skills and social interaction behaviors were less apparent in the non-game simulation, which consists mainly of unguided, task-oriented activities. We speculate that the game-based simulation was effective as it provided the participants with the opportunity for "mindful play" [45] that evoked their inquiry mind and cognitive efforts. The goals set in the game-based simulation was challenging as they require good understanding of the complex relationship between the refraction of light and different medium. While achieving such goals was not easy for most participants, we suspect that the game-based condition provided them with a "pleasantly frustrating feeling", which is one of the essential elements of good game design suggested by Gee [46].

While this study suggests the need to have a certain level of structure in the exhibit design, we also agree with other researchers that have pointed out the danger of formalizing informal learning experiences [47]. As suggested by Allen [3], science centers as a social institution face the "constructivist dilemma" in their institutional goals to meet both the goal to entertain diverse types of visitors and the goal to be an educational space where people can learn about canonical science. Our study may provide some implications regarding how to create exhibit environments where visitors feel entertained with a state of free choice, yet also experience learning some forms of complex science concepts. Our initial experiment with *ARfract* has attempted to tackle the unexplored area of research concerning the design of interactive simulation exhibits, and the sequencing of different types of simulation. Other important questions still remain, such as "what is the optimal level of challenges that best engage visitors?" and "how to better promote collaborative social learning to happen among multiple visitors?". These are questions that we hope to answer in our future research programs.

7 Conclusion and Future Work

This paper presents our proof-of-concept research and the results from an experimental study. In this study, we designed a multi-user hybird AR simulation called *ARfract* where learners can actively explore, speculate, and reflect on their learning as they engage complex science concepts through interacting with an AR simulation. For the two types of AR simulation we designed, namely game-based simulation and non-game based simulation, we examined how the order of the different types of simulation affects learning of abstract science concepts. Overall, the results from the experimental study suggest that the participants who experienced the game-based simulation prior to the non-game simulation gained higher conceptual learning than those who experienced the simulation with the revised order. We also found some important findings related to the design of wearable AR and projection methods from embodied interaction perspectives.

Although we witnessed some promising results in this initial research, we acknowledge that there could be some limitations. First, the sample size was rather small, which may limit the generalizability of findings to other contexts and user groups. We limited the number of participants to 10 dyads as this experiment was out first attempt to test the efficacy of hybrid AR simulation with users. Further given the complexity of the technical set-up and the availability of wearable AR glasses, we concluded that it was better to keep the sample size manageable in this experiment. In future research, we plan to expand our experiment with a higher number of users involved. Second, given the comparable results in the post-test scores between the GF and NF groups, our knowledge test assessment may have had a ceiling effect, limiting our ability to measure participants' learning. Third, while this experiment was conducted in a laboratory setting due to the complexity of technical set-up and the control of variables, it is important for this learning experience to be tested as interactive exhibits with a group of real visitors in a science center setting in future research. Fourth, since the purpose of our study was to examine the impact of sequencing different types of simulation, we did not provide any results that compare the impact between AR-based simulation and other presentation methods. We suggest future studies to explore the efficacy of AR simulation with other comparative conditions such as learning with real light sources and learning with 2-D interactive simulations. The learning efficacy and usability of only game-based simulations or only non-game simulations, not in the context of different presentation order, can be also further examined in future research studies.

In conclusion, the major contribution of this research is proof-of concept and implementation of a new simulation-based learning environment using hybrid AR technologies. The wearable visualization that enables embodied interaction with visual objects exemplifies how the affordances of wearable technologies could support meaningful learning of complex science phenomena in informal learning settings. In this paper, we also suggested some design and technical challenges in hybrid AR technologies that future research need to further examine. AR technologies supporting gesture-based interaction with multiple users, for example, need to be designed to support a strong congruency between action and conception. In future studies, we will explore the pedagogical and technological design of the mixed reality exhibit in context, detail the effect of the design features, and unpack how visitors are engaged with this simulation through the in-depth analysis of observed interaction and discourse.

Acknowledgment

Portions of this research study were presented at the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2015R1C1A1A02037692) and by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the "ICT Consilience Creative Program" (IITP-R0346-16-1007) supervised by the IITP (Institute for Information & communications Technology Promotion)

References

[1] P. Bell, B. V. Lewenstein, S. A. W., and M. A. Feder, *Learning Science in Informal Environments: People, Places, and Pursuits*. National Academies Press, 2009.

[2] P. Sommerauer and O. Müller, "Augmented reality in informal learning environments: A field experiment in a mathematics exhibition," *Comput. Educ.*, vol. 79, no. 2014, pp. 59–68, 2014.

[3] S. Allen, "Designs for learning: Studying science museum exhibits that do more than entertain," *Sci. Educ.*, vol. 88, no. S1, pp. S17–S33, 2004.

[4] N. R. Council, *Learning Science Through Computer Games and Simulations*. Washington, D.C.: National Academies Press, 2011.

[5] D. B. Clark, E. E. Tanner-Smith, and S. S. Killingsworth, "Digital Games, Design, and Learning: A Systematic Review and Meta-Analysis," *Rev. Educ. Res.*, vol. 86, no. 1, pp. 79–122, Mar. 2016.

[6] D. A. Arena and D. L. Schwartz, "Experience and explanation: Using videogames to prepare students for formal instruction in statistics," *J. Sci. Educ. Technol.*, vol. 23, no. 4, pp. 538–548, Aug. 2014.

[7] J. M. Kant, K. Scheiter, and K. Oschatz, "How to sequence video modeling examples and inquiry tasks to foster scientific reasoning," *Learn. Instr.*, 2017.

[8] P. Milgram and F. Kishino, "A Taxonomy of Mixed Reality Visual Displays," *IEICE Trans. Inf. Syst.*, vol. E77–D, no. 12, 1994.

[9] M. Billinghurst, A. Clark, and G. Lee, "A Survey of Augmented Reality," *Found. Trends® Human–Computer Interact.*, vol. 8, no. 2–3, pp. 73–272, Aug.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TLT.2017.2750673, IEEE Transactions on Learning Technologies

2015.

[10] P. Milgram and H. Colquhoun, "A Taxonomy of Real and Virtual World Display Integration," *Mix. Reality-Merging Real Virtual Worlds*, pp. 5–30, 1999.

[11] L. S. Myneni, N. H. Narayanan, S. Rebello, A. Rouinfar, and S. Pumtambekar, "An Interactive and Intelligent Learning System for Physics Education," vol. 6, no. 3, pp. 228–239, 2013.

[12] R. Lindgren and J. M. Moshell, "Supporting children's learning with body-based metaphors in a mixed reality environment," *Proc. 10th Int. Conf. Interact. Des. Child. - IDC'11*, vol. Ann Arbor, pp. 177–180, 2011.

[13] S. Price and Y. Rogers, "Let's get physical: The learning benefits of interacting in digitally augmented physical spaces," *Comput. Educ.*, vol. 43, no. 1–2 SPEC ISS., pp. 137–151, 2004.

[14] M. E. C. Santos, A. Chen, T. Taketomi, G. Yamamoto, J. Miyazaki, and H. Kato, "Augmented Reality Learning Experiences: Survey of Prototype Design and Evaluation," *IEEE Trans. Learn. Technol.*, vol. 7, no. 1, pp. 38–56, 2014.

[15] J. Bacca, S. Baldiris, R. Fabregat, and S. Graf, "Augmented Reality Trends in Education: A Systematic Review of Research and Applications," *Educ. Technol. Soc.*, vol. 17, no. 4, pp. 133–149, 2014.

[16] K. E. Chang, C. T. Chang, H. T. Hou, Y. T. Sung, H. L. Chao, and C. M. Lee, "Development and behavioral pattern analysis of a mobile guide system with augmented reality for painting appreciation instruction in an art museum," *Comput. Educ.*, vol. 71, no. July 2015, pp. 185–197, 2014.

[17] R. Lindgren and M. Johnson-Glenberg, "Emboldened by Embodiment Six Precepts for Research on Embodied Learning and Mixed Reality," *Educ. Res.*, vol. 42, no. 8, pp. 445–452, 2013.

[18] M. Bower and D. Sturman, "What are the educational affordances of wearable technologies?," *Comput. Educ.*, vol. 88, pp. 343–353, 2015.

[19] T. Wu, C. Dameff, and J. Tully, "Integrating Google Glass into simulation-based training: experiences and future directions," *J. Biomed. Graph. Comput.*, vol. 4, no. 2, p. p49, Apr. 2014.

[20] T. Coffman and M. B. Klinger, "Google Glass: Using Wearable Technologies to Enhance Teaching and Learning," *Soc. Inf. Technol. Teach. Educ. Int. Conf.*, vol. 2015, no. 1, pp. 1777–1780, Mar. 2015.

[21] T. C. Huang, C. C. Chen, and Y. W. Chou, "Animating eco-education: To see, feel, and discover in an augmented reality-based experiential learning environment," *Comput. Educ.*, vol. 96, pp. 72–82, 2016.

[22] H.-K. Wu, S. W.-Y. Lee, H.-Y. Chang, and J.-C. Liang, "Current status, opportunities and challenges of augmented reality in education," *Comput. Educ.*, vol. 62, pp. 41–49, Mar. 2013.

[23] L. Bedford, "Storytelling: The Real Work of Museums," *Curator: The Museum Journal*, vol. 44, no. 1. pp. 27–34, 2001.

[24] K. L. Ma, I. Liao, J. Frazier, H. Hauser, and H. N. Kostis, "Scientific storytelling using visualization," *IEEE Comput. Graph. Appl.*, vol. 32, no. 1, pp. 12–19, Jan. 2012.

[25] J. DeWitt and M. Storksdieck, "A Short Review of School Field Trips: Key Findings from the Past and Implications for the Future," *Visit. Stud.*, vol. 11, no. 2, pp. 181–197, 2008.

[26] J. P. Gutwill and S. Allen, "Deepening Students' Scientific Inquiry Skills During a Science Museum Field Trip," *J. Learn. Sci.*, vol. 21, no. 1, pp. 130–181, 2012.

[27] S. A. Yoon, K. Elinich, J. Wang, J. B. Van Schooneveld, and E. Anderson, "Scaffolding informal learning in science museums: How much is too much?," *Sci. Educ.*, vol. 97, no. 6, pp. 848–877, 2013.

[28] D. L. Schwartz and J. D. Bransford, "A Time For Telling," *Cogn. Instr.*, vol. 16, no. 4, pp. 475–5223, 1998.

[29] M. Kapur, "Productive Failure," *Cogn. Instr.*, vol. 26, no. 3, pp. 379–424, Jul. 2008.

[30] M.-B. Ibanez, A. Di-Serio, D. Villaran-Molina, and C. Delgado-Kloos, "Support for Augmented Reality Simulation Systems: The Effects of Scaffolding on Learning Outcomes and Behavior Patterns," *IEEE Trans. Learn. Technol.*, vol. 9, no. 1, pp. 46–56, 2016.

[31] M. S. Birnbaum, N. Kornell, E. L. Bjork, and R. A. Bjork, "Why interleaving enhances inductive learning: The roles of discrimination and retrieval," *Mem. Cognit.*, vol. 41, no. 3, pp. 392–402, 2013.

[32] S. H. K. Kang and H. Pashler, "Learning Painting Styles: Spacing is Advantageous when it Promotes Discriminative Contrast," *Appl. Cogn. Psychol.*, vol. 26, no. 1, pp. 97–103, Jan. 2012.

[33] L. Malinverni and N. Parés, "Learning of Abstract Concepts through Full-Body Interaction: A Systematic Review," *J. Educ. Technol. Soc.*, vol. 17, no. 4, pp. 100–116, 2014.

[34] A. N. Antle, M. Droumeva, and G. Corness, "Playing with the sound maker: do embodied metaphors help children learn?," *Proc. 7th Int. Conf. Interact. Des. Child. - IDC '08*, pp. 178–185, 2008.

[35] M. Bower and D. Sturman, "What are the educational affordances of wearable technologies?," *Comput. Educ.*, vol. 88, pp. 343–353, Oct. 2015.

[36] E. Rosenbaum, E. Klopfer, and J. Perry, "On location learning: Authentic applied science with networked augmented realities," *J. Sci. Educ.*

Technol., vol. 16, no. 1, pp. 31–45, 2007.

[37] H. Tüzün, M. Yilmaz-Soylu, T. Karakuş, Y. Inal, and G. Kizilkaya, "The effects of computer games on primary school students' achievement and motivation in geography learning," *Comput. Educ.*, vol. 52, no. 1, pp. 68–77, 2009.

[38] M. Ebner and A. Holzinger, "Successful implementation of user-centered game based learning in higher education: An example from civil engineering," *Comput. Educ.*, vol. 49, no. 3, pp. 873–890, 2007.

[39] K.-H. Cheng and C.-C. Tsai, "Affordances of Augmented Reality in Science Learning: Suggestions for Future Research," *J. Sci. Educ. Technol.*, vol. 22, no. 4, pp. 449–462, 2013.

[40] J. Martín-Gutiérrez, J. Luís Saorín, M. Contero, M. Alcañiz, D. C. Pérez-López, and M. Ortega, "Design and validation of an augmented book for spatial abilities development in engineering students," *Comput. Graph.*, vol. 34, no. 1, pp. 77–91, Feb. 2010.

[41] H. Benko, E. Ofek, F. Zheng, and A. D. Wilson, "FoveAR," in *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology - UIST '15*, 2015, pp. 129–135.

[42] S. Bakker, A. N. Antle, and E. van den Hoven, "Embodied metaphors in tangible interaction design," *Pers. Ubiquitous Comput.*, vol. 16, no. 4, pp. 433–449, 2012.

[43] K. McLean and Association of Science-Technology Centers., *Planning for people in museum exhibitions*. Association of Science-Technology Centers, 1993.

[44] H.-J. So, J. H. Lee, and S. J. Oh, "Interaction Analysis Between Visitors and Gesture-based Exhibits in Science Centers from Embodied Cognition Perspectives," *KOREA Sci. ART FORUM*, vol. 25, p. 227, Sep. 2016.

[45] I. Yahya, "Mindful Play! or Mindless Learning!- Modes of Exploring Science in Museums," in *Exploring science in museums*, Athlone, 1996, pp. 123–147.

[46] J. P. Gee, *What video games have to teach us about learning and literacy*. Palgrave Macmillan, 2003.

[47] S. A. Yoon, K. Elinich, J. Wang, C. Steinmeier, and S. Tucker, "Using augmented reality and knowledge-building scaffolds to improve learning in a science museum" *Int. J. Comput. Collab. Learn.*, vol. 7, no. 4, pp. 519–541, 2012.

Seung Jae Oh is a doctoral student at the Department of Creative IT Engineering, Pohang University of Science & Technology (POSTECH), Korea. He earned his BSc in Electrical Engineering at POSTECH. His research focuses on exploring the affordances of emerging technologies from human-computer interaction perspectives.

Hyo-Jeong So is an associate professor in the Educational Technology Department, Ewha Womans University, Korea. She earned her Ph.D. degree from Instructional Systems Technology, Indiana University. Previously, she was an associate professor at the National Institute of Education, Nanyang Technological University, Singapore and at the Department of Creative IT Engineering, POSTECH. She was chair of the Classroom, Ubiquitous and Mobile Technologies Enhanced Learning (*CUMTEL*) SIG, the Asia-Pacific Society for Computers in Education (APSCE).

Matthew Gaydos is an Assistant Professor at SUNY - Korea. He earned his PhD from the University of Wisconsin – Madison, USA in Curriculum and Instruction. His work focuses on improving game design for education, government policy, and the use of game-based media in science and civic education contexts.