Play in the Museum: Design and Development of a Game-Based Learning Exhibit for Informal Science Education

Jonathan P. Rowe[1], Eleni V. Lobene[2], Bradford W. Mott[1], James C. Lester[1]

[1]*North Carolina State University, USA*
[2]*Aon Hewitt, USA*

ABSTRACT
Digital games have been found to yield effective and engaging learning experiences across a broad range of subjects. Much of this research has been conducted in laboratory and K-12 classrooms. Recent advances in game technologies are expanding the range of educational contexts where game-based learning environments can be deployed, including informal settings such as museums and science centers. In this article, we describe the design, development, and formative evaluation of FUTURE WORLDS, a prototype game-based exhibit for collaborative explorations of sustainability in science museums. We report findings from a museum pilot study that investigated the influence of visitors' individual differences on learning and engagement. Results indicate that visitors showed significant gains in sustainability knowledge as well as high levels of engagement in a free-choice learning environment with FUTURE WORLDS. These findings point toward the importance of designing game-based learning exhibits that address the distinctive design challenges presented by museum settings.

Keywords: Game-based learning, educational game design, visitor studies, surface computing tables, museum education, pedagogical agents.

INTRODUCTION
Over the past decade, growing evidence has emerged that games are effective learning tools for a broad range of subjects and student populations (Connolly, Boyle, MacArthur, Hainey, & Boyle, 2012; Clark, Tanner-Smith, & Killingsworth, 2016; Wouters, van Nimwegen, van Oostendorp, van der Spek, 2013). Much of the work has focused on two categories of game-based learning: 1) games for formal education settings such as schools (Habgood & Ainsworth, 2011; Wouters et al., 2013), and 2) serious games, which typically investigate game technologies for training (Johnson, 2010) or increasing awareness of social, geopolitical, or economic issues (Mitgutsch & Alvarado, 2012). While important, these research directions do not address a notable class of educational contexts that stands to benefit as much, or perhaps even more, from the introduction of well-designed educational games: informal education settings, such as science museums.

Formal education settings, and particularly schools, differ from science museums in several important respects (National Research Council, 2009). Schools and museums serve different populations: Schools are responsible for educating school-age children, whereas museums are visited by learners of all ages. Museums have a particular emphasis on affective outcomes, such as sparking interest and excitement, and they rarely use tests or grades. In schools, learners typically sit at designated desks for prescribed class periods. Museums are free-choice learning environments; visitors come and go as they please, and there is no teacher to provide instruction. Museum learning, quite literally, *looks different* from classroom learning. These distinctions have significant implications for the design of educational games—including pedagogical and gameplay designs—and they merit careful consideration to ensure effective and engaging learning experiences.

In this article, we explore the design of game-based learning environments for museums by presenting a case study of FUTURE WORLDS, a game-based learning environment for collaborative explorations of environmental sustainability. FUTURE WORLDS is an interactive exhibit designed to enrich museum visitors' understanding of sustainability, promote collaboration and scientific reasoning, and foster engagement in environmental science. We outline the design principles that drove the iterative development of FUTURE WORLDS, and we illustrate the methods utilized to investigate FUTURE WORLDS' effectiveness for fostering learning and engagement.

BACKGROUND AND RELATED WORK

The affordances of digital games align naturally with the goals of museum education, such as fostering engagement in science, and enabling learners to manipulate, test, and explore hypotheses about the natural world (National Research Council, 2009). However, designing game-based learning environments presents several challenges. Game design is multidisciplinary, requiring close collaboration between software developers, educators, artists, testers, and other specialists. Games are complex, requiring myriad design decisions with uncertain impacts on learners' experiences. Most notably, there is a dearth of evidence-based research on the design principles and methods necessary for creating effective game-based learning environments.

Reviews of the game-based learning literature have broadly concluded that games can yield positive learning outcomes across a range of educational subjects (Connolly et al., 2012). In recent years, a pair of prominent meta-analyses independently concluded that, in general, digital game technologies are often more effective than traditional instructional methods in fostering learning and retention (Clark et al., 2016; Wouters et al., 2013). Expanding on this conclusion, Wouters et al. (2013) advise, "the next step is more value-added research on specific game features that determine ... effectiveness" (p. 262). Clark et al. (2016) echo this argument, concluding, "[Research on game-based learning] should thus shift emphasis … to cognitive-consequences and value-added studies exploring how theoretically driven design decisions influence situated learning outcomes" (p. 116).

With the aim of identifying evidence-based design principles for game-based learning, Isbister, Flanagan, and Hash (2010) conducted interviews with experienced game developers to identify key design practices used within the professional game industry. The interviewees described themes such as emphasizing fun as a central design value, requiring high levels of polish and well-tuned end-user experiences, emphasizing deep learning content rather than "bolted on" learning materials, supporting collaboration and specialization, designing for role-playing and emotional engagement, and including affordances for exploring complex systems. While the identified themes are high-level and abstract, they do describe characteristics often lacking in game-based learning environments. Isbister et al. (2010) account for this by pointing to problems in the communication pathways between practitioners and theorists, describing situations where design choices are not well articulated or discussed critically.

In other work, Linehan, Kirman, Lawson, and Chan (2011) describe methods for educational game design rooted in applied behavior analysis. Their framework emphasizes personalized instruction and mastery learning, describing a four-step design pattern that involves 1) defining and measuring learners' behavior, 2) recording and analyzing changes in learners' behavior, 3) providing immediate corrective feedback, and 4) tailoring game events based on learners' performance. It is notable that the guidelines from Linehan et al. are theoretically grounded, but unfortunately they are not examined in the context of an actual educational game.

A handful of empirical studies have been conducted to investigate the learning impacts of specific game design decisions. Habgood and Ainsworth (2011) found that tight integration of subject matter and game mechanics in a game for elementary mathematics yielded enhanced learning outcomes compared to versions that separated content and gameplay. Other work has investigated the impacts of narratives in games (Adams, Mayer, MacNamara, Koenig, & Wainess, 2012; Rowe, McQuiggan, Robison, & Lester, 2009). Results suggest that narratives may be effective in fostering student motivation, but narrative designs should be crafted sparingly in order to avoid distracting learners or imposing extraneous cognitive load.

Much of this research has taken place in the context of formal education settings, such as K-12 classrooms. However, the design requirements for game-based learning in museums differ from games in schools because of the distinct characteristics of museum-based learning. As outlined in a report from the National Research

Council (2009), the informal science education community has identified six interrelated strands of informal science learning: sparking interest and excitement in science; enabling learners to understand, remember, generate, and use science concepts; encouraging learners to test, explore, question, and observe the natural and physical world; fostering reflection on the scientific process; creating opportunities for learners to participate in scientific activities; and encouraging learners to think of themselves as scientists.

Museums and schools share many of these objectives, but they prioritize the objectives differently. Schools are responsible for preparing students to meet academic standards focused on student knowledge, skills, and abilities. Attendance is often compulsory, and instructional practices are aligned with school objectives. In contrast, the voluntary nature of many museum visits leads toward an increased emphasis on non-cognitive outcomes, such as visitor interest and engagement. Museum visitors typically spend brief amounts of time at each exhibit—a few minutes or shorter—which limit expectations for deep knowledge gains. Museums also have different educational resources (e.g., panoramic theaters, fossil collections) and physical layouts (e.g., open floor plans) than schools, which shape the learning experiences that are available to visitors. Discerning the extent to which a game-based learning environment embodies the six strands of informal science learning, as well as the extent to which it accounts for the distinct capabilities and constraints of museum settings, provides a framework for evaluating a game-based learning exhibit's design.

Recent years have witnessed growing interest in the application of digital game technologies in informal settings, including animated pedagogical agents (Lane et al. 2013; Schulman, Sharma, & Bickmore, 2008), interactive tabletop-based games (Horn et al., 2012), and mobile games (Greenspan & Whitson, 2013). For example, advanced learning technologies in museums now utilize the high-end rendering capabilities of game engines so that virtual humans serve as simulated docents. These docents offer personalized guidance and feedback to visitors through natural language and affective expression (Lane et al., 2013). Game-based learning exhibits have also been developed for a range of subjects, including evolution (Horn et al., 2012), sustainability (Antle, Bevans, Tanenbaum, Seaborn, & Wang, 2011), and history (Greenspan & Whitson, 2013). These exhibits have shown promise for promoting engagement and visitor enthusiasm by leveraging game designs centered on brief visitor interactions, multi-touch interfaces, and multi-user hardware technologies (Antle et al., 2011; Horn et al., 2012;). However, there has been relatively little empirical research on visitor learning and individual differences with game-based exhibits in science museums. Further, few projects have explicitly examined the design factors that distinguish game-based learning in museums from game-based learning in schools. In the next section, we outline these design factors, which shaped the development of the FUTURE WORLDS exhibit.

DESIGN FACTORS FOR GAME-BASED LEARNING IN MUSEUMS

During the early stages of the project, we held a series of discussions with partners at the North Carolina Museum of Natural Sciences about interactive science exhibits and the dynamics of visitor learning. Further, we reviewed several interactive exhibits in prominent science museums across the eastern United States to inform our efforts to gather software requirements for the game's design. From this analysis, we identified five overarching design principles for FUTURE WORLDS to address the objectives, affordances, and constraints of game-based learning in museums (Table 1).

1. **Low barrier to entry**. Because interactions with museum exhibits can be either short or long in duration, embedding game mechanics that are accessible and quick-to-learn is critical. For this reason, we adopted natural user interfaces—in our case multi-touch interaction—and highly streamlined heads-up displays (HUD) to guide learner interactions. Similarly, we sought to leverage learners' prior knowledge whenever possible, adopting multi-touch controls that are widely adopted in modern computing (e.g., tap, swipe, pinch-to-zoom, two-finger rotation), as well as game paradigms that are widely recognized by the public (e.g., SimCity-style visual appearance).
2. **Exploration and curiosity**. Effective interactive exhibits typically avoid didactic instructional methods, such as lengthy text presentations and non-interactive animations. Our design, therefore, emphasizes

Table 1. Alignment between museum-based learning attributes and game design principles.

Museum-Based Learning Attributes	Design Principles for Game-Based Exhibits
Emphasize interest and excitement	• Foster exploration and curiosity • Present melodramatic feedback • Leverage novel hardware
Feature short dwell times	• Prioritize low barriers to entry
Appeal to diverse learners (i.e., age, gender)	• Prioritize low barriers to entry • Utilize inviting aesthetics with broad appeal
Support solo and collaborative interactions	• Leverage novel hardware • Provide flexible user experiences
Support active and passive engagement	• Leverage novel hardware • Provide flexible user experiences
Account for free-choice movement between exhibits	• Foster exploration and curiosity • Present melodramatic feedback

learners' direct manipulation of the virtual environment, as well as explorations of cause-and-effect and interrelations between disparate environmental elements. We sought to minimize the amount of text in the game and avoid direct instruction about environmental science concepts.

3. **Immediate and dramatic feedback**. Immediate feedback is important for any educational setting, but discussions with museum partners yielded recommendations for incorporating melodramatic feedback (in their words, "use fireworks") to convey the effects of learners' decisions. This advice was specifically targeted at young learners, with the implication that subtle effects are often missed or ignored in museum-based learning. Museum spaces often house many competing exhibits, and learner attention is easily diverted. Therefore, we sought to integrate rich combinations of animation, color, and sound to make the effects of young learners' actions clear.
4. **Inviting visual aesthetics with broad appeal**. Although the project targeted young learners, parents and guardians often accompany these visitors. Given the broad range of potential participants found in science museums, we sought to devise a visual style that was inviting and had broad family appeal; we chose not to adopt a visual style that was edgy or dystopian in conveying the impacts of environmental decisions. While labor-intensive, we also sought to emphasize visual "polish," consistent with recommendations from the educational game design literature (Isbister et al., 2010). The importance of visual polish was particularly salient in our review of exhibits in prominent science museums. High-quality visual presentations are the standard, not the exception, in museum-based learning. This emphasis on visual quality is a notable differentiator from many game-based learning environments for classroom learning.
5. **Novel hardware platforms**. Given the leisure-based nature of many museum visits, and relatively high bar for learners' attention[1], we determined that using a novel hardware platform was an important opportunity for fostering learner interest and engagement. We decided to utilize surface computing hardware to create interactive learning experiences that were not feasible in most school or home settings. This criterion is contrasted with formal education settings, where even rudimentary games are often viewed as being more engaging than conventional instructional methods, such as lectures. Surface computing tables simultaneously integrate multi-touch interfaces, which are immediately familiar to

[1] In our partner museum, competing exhibits include panoramic theaters, dinosaur skeletons, live animals, and open labs where the public interact with real scientists.

Figure 1. Early artist rendering of FUTURE WORLDS multi-display exhibit

many learners, with novel multi-user collaboration possibilities. We then extended the platform by integrating the surface computing table with a second, vertically mounted high-definition digital display, doubling the available screen real estate for the exhibit (Figure 1). One collaborator described the platform as akin to a "giant Nintendo DS."

In addition to these principles, we came to recognize several interaction patterns that distinguish museum-based learning from other settings, particularly classroom learning. These point to the need for a sixth design principle: providing flexible user experiences that accommodate the broad range of visitor interactions typified by museum-based learning.

1. **Solo vs. collaborative interactions**. Learners approach museum exhibits in many different types of configurations. Learners may explore museums individually, engaging in solo learning experiences involving limited conversation and social interaction. Young learners may explore exhibits while accompanied by siblings, parents or guardians. Adults, who may or may not directly interact with the exhibit themselves, will often scaffold young learner's interactions by asking questions or giving advice. In another interaction pattern, learners may approach exhibits that are already partially occupied, engaging in collaborative or cooperative learning with "strangers" with whom they share no prior relationship. Alternatively, learners may approach exhibits as part of large field trip groups, exploring exhibits collaboratively with large groups of peers with whom they have strong social bonds. Given these distinct types of configurations, game-based learning designs for museums should be compatible with a broad range of solo and multi-player interaction paradigms.
2. **Short vs. long durations**. Learner interactions may be shallow and brief—5 second interactions that are quickly abandoned—or they may be deep and extensive—30 minute explorations involving detailed conversations and reflection. Paradoxically, learner interactions may also be long in duration but shallow in terms of learning. In these cases, learners play with surface features of the exhibit but fail to deeply engage with the subject matter. Gameplay and content in a museum-centric game-based learning environment should be devised to accommodate these different extremes of engagement, providing positive educational value in as many scenarios as possible.
3. **Active vs. passive engagement**. Active prolonged engagement is an oft-cited goal in museum exhibit design (Horn et al., 2012). However, in practice, learners may engage with an exhibit in many different ways. Some learners confidently approach an exhibit and begin physically interacting immediately. Others initially hold back, passively watching or deliberating whether to explore the exhibit more

actively. Learners may also passingly engage with an exhibit while walking en route to another destination, judging whether to explore the exhibit at a later point in time. A salient indicator of these differences in engagement is physical proximity; active engagers are often physically close to the exhibit, while passive engagers are often farther away. Another important indicator is dwell time, which denotes the amount of time that a visitor spends at an exhibit. In our work, we attend to both factors together, investigating how long visitors spend at several different tiers of proximity to the exhibit. Given individual differences in engagement, game-based learning exhibits should be designed to support learning across different physical distances and durations, gradually easing learners toward higher levels of active engagement as they are comfortable.

Having identified these design principles and museum-centric interaction patterns, we embarked on the design and development of a game-based learning environment for sustainability education in museums, FUTURE WORLDS.

DESIGN AND DEVELOPMENT OF FUTURE WORLDS

To create FUTURE WORLDS, we adopted an iterative design process that involved close interdisciplinary collaboration between developers (game designers, software engineers, digital artists), educators (informal science educators, elementary science educators), and subject matter experts (environmental engineers, environmental scientists). The game design model involved several iterations of designing, developing, evaluating, and refining prototypes of FUTURE WORLDS, including both paper and software-based prototypes.

FUTURE WORLDS

FUTURE WORLDS is a prototype game-based learning exhibit about environmental sustainability for children ages 9–12. The exhibit integrates turn-based strategy game mechanics, pedagogical agents, and surface computing tables to support collaborative explorations of environmental sustainability. In FUTURE WORLDS, learners solve sustainability-centered problems by investigating the impacts of alternate environmental decisions in a 3D simulated environment. Learners explore environmental decisions—such as modifying a region’s electricity portfolio or a farm’s waste management practices—through collaborative, multi-touch interactions on the exhibit’s interactive tabletop display. The effects of learners’ environmental decisions are realized in real-time through rich 3D virtual environments, and they are accompanied by narrated explanations from a virtual docent who observes the learners’ actions.

The design of FUTURE WORLDS is intentionally aligned with several strands of informal science learning (National Research Council, 2009). Specifically, FUTURE WORLDS offers opportunities for enriching visitors’ understanding of environmental sustainability; it fosters high levels of visitor engagement; it enables visitors to test, manipulate, and observe changes in the simulated environment; and it fosters collaboration and conversation about environmental science. FUTURE WORLDS also accounts for the distinctive capabilities and constraints of science museums. The game accommodates different levels of visitor engagement, ranging from brief passive encounters to extended hands-on experiences. FUTURE WORLDS’ multi-touch gameplay is designed to be easily accessible, particularly for learners with minimal game-playing experience. Further, the game’s 3D graphics are designed to appeal to a broad range of visitors.

The prototype exhibit’s physical design is comprised of two adjacent digital displays: a horizontally oriented Samsung SUR40 interactive tabletop, and a vertically oriented non-interactive 50” high-definition television (Figure 1). The configuration provides an integrated two-screen setup; visitors congregate around the horizontal display to explore the science simulation through multi-touch interactions, and the vertical display provides additional screen real estate for explanations of sustainability concepts, which are also accessible to learners standing farther away from the exhibit.

FUTURE WORLDS focuses on three integrated themes of sustainability: water, food, and energy. Visitors’ objective during learning interactions with FUTURE WORLDS is to use the interactive tabletop display to collaboratively (or individually) reconfigure an unsustainable virtual environment into a sustainable environment. Learners can engage in deep, extended interactions with FUTURE WORLDS —solving sustainability problem scenarios and identifying complex relationships between alternate environmental decisions—or they can engage in shallow interactions with FUTURE WORLDS, tapping on the virtual environment and briefly listening to a narrated explanation before moving on.

As an illustration of a typical learning interaction with FUTURE WORLDS, consider the following scenario. A small group of learners approaches the exhibit. Tapping on the start screen causes a 3D model of Earth to rotate into view. From the globe, a small environmental region from the United States' eastern coast emerges. The virtual environment represents the first problem-solving scenario: a portion of a simulated watershed that is in an unsustainable state.

The virtual environment is divided into several discrete, hexagon-shaped locations (Figure 2). Each hexagon encompasses an atomic geographic region that can be acted upon by learners as part of solving the environmental sustainability problem scenario. The virtual environment's three-dimensional appearance is stylized, and at first it appears to be brown and lifeless. The environment's appearance is a melodramatic visualization illustrating that the virtual environment is "unhealthy," that is, the environment is in an unsustainable state. Shortly afterward, a short description of the scenario's sustainability-centric objective is presented. The scenario focuses on issues related to farming, particularly waste management and electricity sources. Learners are instructed to explore alternate choices for improving the sustainability of the farms. At this stage, learners are free to begin manipulating the virtual environment to explore the effects of alternate environmental decisions.

Learners can manipulate the virtual environment using multi-touch gestures, such as pinch-to-zoom and rotations. Additionally, learners can tap on various hexagon-shaped locations in the virtual environment, request information about the different factors that bear on the environment's sustainability, and change the behavior of the farms. For example, one of the farms appears to be contaminating a nearby river through fertilizer runoff from a crop field. A learner taps on the farm, and a menu appears offering several options for changing the farm's fertilizer runoff management practices. In order to learn more about the farm's current management practices, the learner taps an "Info" button on the menu. In response, a virtual docent appears on the second vertical display providing a short narration about fertilizer runoff and waste management.

Another example of the use of FUTURE WORLDS is that a learner encounters the option to add a riparian buffer to the hog farm. The docent explains that a riparian buffer consists of plants and tree roots that can absorb hazardous nitrogen levels, which leach from synthetic fertilizers used by many farms into nearby water bodies. The learner decides that installing a riparian buffer is worth exploring, and she taps on the menu to enact it. In response, the virtual environment begins to change in appearance: the farm's 3D models transform as the brown-colored fertilizer runoff disappears and a stretch of small bushes, shrubs, and tall grass appears between the farm and adjacent riverbank (Figure 2). Nearby terrain grows greener, and portions of the environment begin to animate. The 3D virtual environment—replete with aesthetic lighting, detailed 3D models, high-resolution textures, and believable animations—provides a dynamic visualization of the learners' environmental actions through immediate feedback, in effect making learners' sustainability choices "come to life." Using this feedback, learners explore the sustainability simulation and solve the problem scenario (Figure 2).

Figure 2. Successive stages of a virtual environment in FUTURE WORLDS.

Paper Prototyping

After identifying design requirements to guide the development of FUTURE WORLDS, we undertook an iterative paper prototyping process that involved developing, testing, and refining several paper-based versions of the system (Figure 3). Paper prototypes emulate the visual appearance and interaction design of the final game environment while providing a low-cost method for rapidly creating, deploying, and refining alternate interaction designs. Paper prototypes resemble a “board game” version of the final system—they are typically made using paper and cardboard cutouts—and they source difficult-to-model aspects of the system’s functionality to a human “game master.” Pilot testing paper prototypes with young learners shares similarities with Wizard of Oz studies (Kelly, 1983). In both cases a human emulates functionality that will eventually be provided by software, such as simulation systems or artificial intelligence functionalities.

Figure 3. FUTURE WORLDS paper prototype.

During the project’s first phase, the team conducted two studies to test early paper prototypes of FUTURE WORLDS. The first study was held during a teen-focused event at the North Carolina Museum of Art, and it focused on FUTURE WORLDS’ interaction design and visual style. The second study was held at the North Carolina Museum of Natural Sciences, and it focused on a refined version of the FUTURE WORLDS paper prototype, investigating the complexity of the environmental problem-solving tasks.

During the first study, 11 learners ranging in age from 13-17 interacted with the paper prototype. The prototype focused almost entirely on tradeoffs between energy generation and pollution. Its main purpose was to support investigations of FUTURE WORLDS’ core gameplay dynamics. In addition to testing basic interaction design, the study provided an opportunity to investigate methods for presenting scientific information, and compare alternate visual styles for the virtual agents. Learners worked individually and collaboratively to solve two problem scenarios with assistance from a project team member. All learners solved the scenarios, and several learners identified alternative solutions to the scenarios. However, the research team observed that learners did not appear to engage deeply with the descriptions of alternate energy sources or consider a broad range of environmental factors in their decision making. The project team reviewed the paper prototype following this first focus group in order to generate feedback about the science content and presentation. The paper prototype was modified in order to enhance the science content, refine the task descriptions, introduce an additional problem scenario, and modify the input mechanics.

The second study was held in coordination with a science summer camp program hosted by the North Carolina Museum of Natural Sciences. The study was designed to examine the revised implementation of the FUTURE WORLDS paper prototype. The focus group involved 17 learners in total. The learners were divided into groups so that learners could work on the paper prototypes in pairs while a designated researcher oversaw each paper prototype. Each group interacted with the paper prototypes for approximately 15–20 minutes.

During the focus group’s pre- and post-assessments, learners completed a drawing task that involved creating a picture of what is an environment based on their understanding. Participants were asked to label parts of the drawing and write an explanation of what makes their drawing a representation of an environment. Learners produced images of forests, towns, water bodies, animals, and the sky. As part of the post-assessment that was administered after the FUTURE WORLDS experience, learners were given the opportunity revise their original drawings. Of those learners, 16 elected to make changes and improve their representation of the environment. All of these learners chose to make additions and not remove aspects of their original drawings. The most common new inclusions were representations of energy sources such as wind (windmills), solar (the sun), petroleum (oil power plants), hydroelectric power (water sources) and human beings, providing preliminary evidence of positive gains in sustainability understanding.

Iterative Development

Although agile-based software development processes are common practice in the game industry, only in recent years have iterative development processes begun to gain attention in the research literature on educational game design (Chamberlin, Trespalacios, & Gallagher, 2012). Our process further contributes to this growth: Upon establishing a benchmark design through several iterations of paper prototyping and pilot tests, we began software production on the FUTURE WORLDS game-based learning environment. FUTURE WORLDS was built with the Unity game engine. An agile-based development process was employed, which involved identifying, prioritizing, and developing targeted sets of features for 2-4 week development "sprints." During each sprint, software development and art production activities focused only on the features identified for that sprint. At the conclusion of each sprint, an incremental build of the system was demoed internally to the project team, as well as outside collaborators, for constructive feedback. Software production proceeded in this manner for approximately six months.

MUSEUM PILOT TEST

A study was conducted at the North Carolina Museum of Natural Sciences to investigate visitors' learning and engagement with the FUTURE WORLDS prototype game-based learning exhibit. Participants were summer campers enrolled in existing summer camp programs at the museum and visited the exhibit in small groups. The study was designed to investigate the prototype exhibit's impacts in a free-choice learning environment that approximated real-world use—this included several competing co-located exhibits—as well as the influence of learners' individual characteristics on sustainability learning and dwell time.

Participants

A total of 43 participants interacted with FUTURE WORLDS during the study. The campers were grouped into separate cohorts divided across three sessions (*N*=14, 16, 13). Of the sample of 43 campers, a majority was male (57.4%). Participants ranged from ages 8-14 with a mean age of 10.7 years. The majority of participants identified as Caucasian (75.9%), followed by Asian/Pacific Islander (7.4%), Black (3.7%), and American Indian/Alaskan Native, Hispanic, and Other (1.9%).

Materials and Apparatus

Learners completed three complementary measures of sustainability concept understanding: a personal meaning map, a sustainability identification task, and a sustainability image-sorting task. These measures were devised to assess visitors' ability to understand and remember sustainability concepts, while remaining consistent with a spirit of voluntary participation and leisure-based learning in museums. To investigate learner engagement, video recordings of participant behavior were recorded and analyzed to determine dwell times.

Personal Meaning Maps. In order to investigate learners' conceptual understanding of sustainability, participants completed personal meaning maps (PMMs). Each PMM consisted of a blank piece of paper with a brief set of instructions and a prompt phrase: sustainability. Participants used a blue ink pen to write or draw words, phrases, and pictures about what they thought and knew about the prompt word on the blank paper. After leaving the exhibit space, participants were given the opportunity to revise their PMM using a red ink pen. After the study, two raters scored each PMM based on the relevance and accuracy of each element included on the page. Typical features that were considered *sustainability-relevant* included drawings of windmills, trees, and solar panels, as well as descriptive words such as “renewable energy” and “ecosystem.” Irrelevant features included drawings of pizza and other foods—some learners appeared to confuse “sustenance” with “sustainability”—as well as words that indicated misconceptions or guessing, such as “strong.” The raters summed the total number of relevant and correct items listed and subtracted the total number of irrelevant and incorrect items included. The inter-rater reliability for the pre-test (r=.84) and post-test (r=.88) grades achieved acceptable levels.

Identification Task. Learners were instructed to inspect an illustrated picture of an environment—which depicted both sustainable and unsustainable environmental practices—and annotate the picture by circling “good” environmental practices and crossing out “bad” practices. Learners’ pictures were returned during the post-test, and participants revised their annotations. Two independent raters scored learners’ annotations using a rubric vetted by subject matter experts. The raters obtained high scoring agreement on both the pre-test (r=.97) and post-test (r=.95) versions.

Image Sorting Task. Learners were given paper copies of ten images depicting various environmental practices (e.g., recycling, solar panels, heavy traffic congestion, off-shore drilling). Participants were asked to organize the images into two categories of their choosing and were instructed to select the two categories such that they contained as similar a number of images as possible. An expert-based categorization of the images into “sustainable” and “unsustainable” categories was considered the most disciplinarily sophisticated response, and this benchmark was used to grade students’ responses. This activity was completed as part of both the pre- and post-test.

Dwell Times. Sessions were video recorded by two cameras strategically placed in the study room to collect data on visitor dwell behavior. Through post-study analyses of the video data, dwell times were determined for each participant in order to assess durations of learner engagement. For each learner, two coders determined total dwell time (time spent interacting with the exhibit to any extent) as well as time spent in each of three possible “tiers of proximity” relative to the FUTURE WORLDS exhibit. The first tier is defined as touching the table or standing directly against the interactive tabletop. The second tier is defined as standing immediately behind the first tier participants. The third tier is one layer behind the second tier, within an approximately 5-foot radius of the exhibit. Inter-rater reliability was established with a subset of the study data and then the remainder was coded independently (k=.70).

Study Design and Procedure

Pre-test measures were administered on the first day of each cohort’s week-long summer camp. The pre-test included the PMM, identification task, sorting task, and demographic questionnaire. Later in the week, learners spent time exploring various parts of the museum, including a room with a thematic focus of “citizen science.” This area served as the study room. The FUTURE WORLDS exhibit (Figure 4) was kept on-site and made available for use by the summer campers when exploring the citizen science area. During each study session, a human docent stood next to FUTURE WORLDS to resolve technical issues and answer questions. Conversation between the docent and participants was kept to a minimum. As part of their summer camp experience, learners were allowed to spend up to 40 minutes in the study room. For each cohort, all participants entered the room at the same time. Learners were told that they could leave the space at any time, and were free to explore the space as they saw fit. Once they left the room, they were not permitted to return.

Including FUTURE WORLDS, there were 13 exhibits in the study room. No other museum visitors had access to the citizen science area during the study. Eleven of the exhibits in the study room were permanent exhibits at the museum. In addition to FUTURE WORLDS, one temporary exhibit was added to serve as a control. This temporary exhibit was the only other exhibit with a human docent, and it consisted of a white board with a sign asking learners to write down the most interesting thing they learned in the citizen science area. Of the 12 distractor exhibits within the space, half of them (N=6) were interactive and more than half were digital (N=7). It should be noted that none of the content in the distractor exhibits overlapped with the content included in FUTURE WORLDS.

A post-test was conducted immediately following each participant’s exit of the citizen science area. It included the same sustainability learning assessments as the pre-test.

Figure 4. Visitors exploring the prototype FUTURE WORLDS game-based learning exhibit.

FINDINGS

To investigate the learning impact of the FUTURE WORLDS exhibit, statistical analyses of the pre- and post-test measures, as well as coded video recordings, were conducted. Paired t-tests indicated that learners showed significant gains in PMM score from pre-test (*M*=0.8, *SD*=1.8) to post-test (*M*=1.2, *SD*=2.3), $t(37) = 2.5$, $p<.05$. There were also significant increases on the identification task from pre-test (*M*=6.0, *SD*=2.5) to post-test (*M*=6.4, *SD*=2.6), $t(37)=3.3$, $p<.05$, as well as significant gains on the image sorting task from pre-test (*M*=7.1, *SD*=3.8) to post-test (*M*=8.7, *SD*=2.7), $t(37)=2.6$, $p<.05$. Correlation analyses were conducted to investigate whether learners' individual characteristics—including age and gender—showed significant relationships with learning outcomes, but none were observed.

For each learner, total dwell time, as well as time spent in each of three proximity tiers, was determined using video recording data. Across all participants, the average dwell time at FUTURE WORLDS was 12.5 minutes. This is promising, given dwell times typically reported in other informal contexts, such as 5.0 minutes (Horn, Solovey, Crouser, & Jacob, 2009) and 4.9 minutes (Lane et al., 2013). However, it should be noted that FUTURE WORLDS dwell times were recorded in a semi-controlled study setting, whereas the above cited dwell times were recorded from observations of the general public in non-controlled settings.

We next investigated the effects of individual differences—specifically gender and age—on dwell time. A two-tailed independent samples t-test revealed a significant effect of gender on dwell time, where girls (*M*=8m:46s, *SD*=5m:21s) spent roughly half the time as boys (*M*=16m:23s, *SD*=10m:33s) engaging with FUTURE WORLDS, $t(35)=2.9$, $p<.05$. To follow up on this analysis, tier-specific dwell time was examined by gender. Results indicated that males (*M*=14m:12s, *SD*=7m:34s) spent significantly more time in the first tier than females (*M*=8m:56s, *SD*=3m:56s), which is the tier of closest proximity to the exhibit, $t(29)=2.3$, $p<.05$.

Additional findings about the influence of learners' individual characteristics emerged from analyses of engagement based on learner age. Results indicated that younger children spent more time at the FUTURE WORLDS exhibit than older children, $t(44)=3.5$, $p<.01$. In fact, children under age 10 spent twice as much time (*M*=20m:50s, *SD*=12m:33s) as children age 10 and older (*M*=9m:56s, *SD*=6m:45s). Observations of mean dwell time by age are shown in Figure 5.

DISCUSSION

Our results suggest that learners' sustainability understanding improves from interactions with FUTURE WORLDS. Furthermore, evidence of extended dwell time, compared to the existing literature, suggests that learners are highly engaged with the exhibit. In combination, these two sets of findings suggest that learner

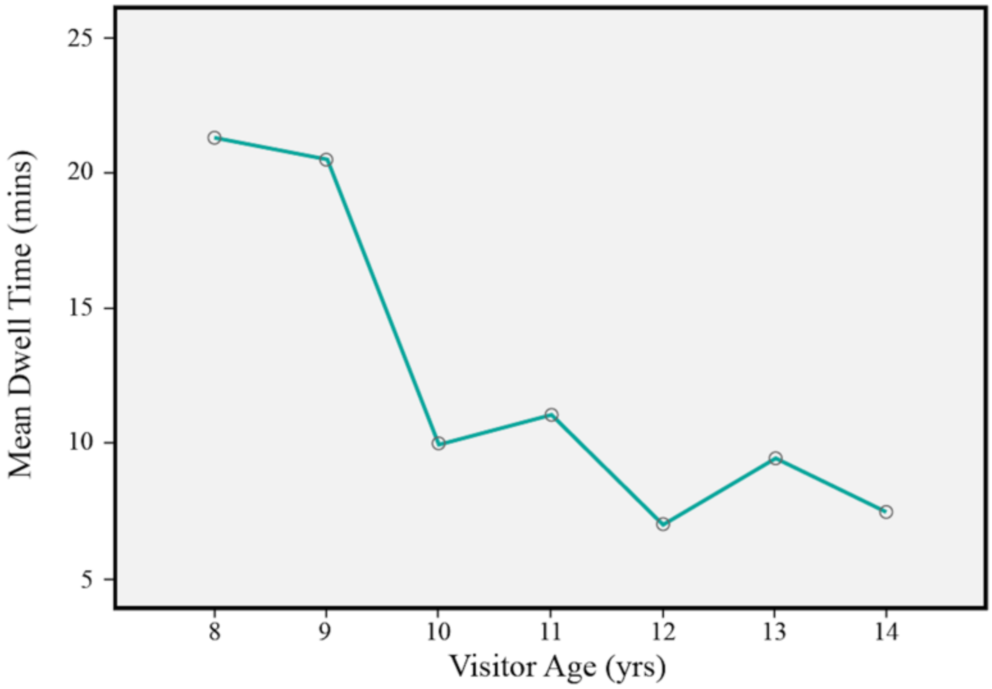

Figure 5. Mean dwell time (in minutes) with FUTURE WORLDS by visitor age (in years).

engagement with FUTURE WORLDS is not superficial; learners are actively engaged for prolonged periods at sufficient depth to yield significant learning gains across three distinct measures of sustainability knowledge.

Our observation that gender and age have significant effects on dwell time point toward engagement-centric design implications for future iterations of FUTURE WORLDS. Regarding gender, several possible mediating factors could explain why girls spent less time with FUTURE WORLDS than boys, such as video game interest or affinity for the game's visual aesthetic style. Gender is widely recognized for its importance in shaping the design and use of entertainment-focused games (Jenson & de Castell, 2010). Our findings support that gender is likely to be similarly important in the design and use of game-based learning exhibits in museums. Additional studies are necessary to isolate what design factors are responsible for the observed gender differences in visitor experiences with FUTURE WORLDS and to determine how FUTURE WORLDS' design should be augmented to appeal to females to an extent that is comparable to males.

Regarding age, results suggest that future iterations of FUTURE WORLDS should incorporate problem-solving scenarios that span a broader range of content and complexity. It is possible that the implemented problem scenarios in the FUTURE WORLDS prototype were sufficiently challenging for young children but were not difficult enough for older children, and thus did not sustain their engagement for extended periods. Older visitors may have solved all of the game's embedded sustainability challenges—in other words, they may have *finished* the game—more quickly than younger learners, leading to shorter dwell times. Incorporating intelligent tutoring system capabilities to dynamically adapt the difficulty of problem scenarios to individual learners, or groups of learners, is a promising way to match scenarios' content complexity to the capabilities of diverse visitors. Adaptive pedagogical planning will require models for automatically detecting learners' individual characteristics as they approach and use exhibits, since administering lengthy pre-tests is not a viable design choice for most informal settings.

Notably, we did not find a relationship between dwell time and learning. We expect that adding curricular material beyond the prototype's current proof-of-concept scope—creating opportunities for more variance in content exposure—could reveal relationships between dwell time and learning that have not thus far been observed.

CONCLUSION AND FUTURE WORK

Game-based learning environments show considerable promise for informal settings such as museums and science centers. Creating game-based learning environments for science museums raises challenges that are distinct from those raised by games for classrooms. Museums and classrooms emphasize different aspects of learning, serve different learner populations, and afford different levels of learner autonomy. To address these challenges, we have identified several design factors for creating game-based learning exhibits in science museums: (1) providing a low barrier to entry, (2) fostering exploration and curiosity, (3) delivering immediate and dramatic feedback, (4) prioritizing visual aesthetics with broad appeal, (5) leveraging novel hardware platforms, and (6) providing flexible user experiences. To investigate these design factors, we pilot tested a game-based learning exhibit called FUTURE WORLDS, which was created using these design criteria as a guide. The pilot study found that visitors achieve significant gains in sustainability knowledge, as well as high levels of engagement, while learning with the game. Boys were observed to actively engage with FUTURE WORLDS for significantly longer durations than girls, and young children engaged with the exhibit for longer periods than did older children. These individual differences underscore the importance of future work investigating adaptive pedagogical functionalities, as can be provided by intelligent tutoring system techniques, as well as automated detectors for diagnosing learners' individual and group characteristics.

ACKNOWLEDGMENT

The authors would like to thank the staff and visitors of the North Carolina Museum of Natural Sciences as well as the IntelliMedia Group at North Carolina State University. This research was supported by the National Science Foundation under Grant DRL-1114655. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Adams, D., Mayer, R., MacNamara, A., Koenig, A., & Wainess, R. (2012). Narrative games for learning: Testing the discovery and narrative hypotheses. *Journal of Educational Psychology*, *104*(1), 235–249.

Antle, A., Bevans, A., Tanenbaum, J., Seaborn, K. & Wang, S. (2011). Futura: Design for collaborative learning and game play on a multi-touch digital tabletop. In M. Gross, N. Nunes, E. Do, S. Brewster, & I. Oakley (Eds.), *Proceedings of the Fifth International Conference on Tangible, Embedded, and Embodied Interaction* (pp. 93–100). New York, NY: ACM.

Chamberlin, B., Trespalacios, J. & Gallagher, R. (2012). The Learning Games Design Model: Immersion, Collaboration, and Outcomes-Driven Development. *International Journal of Game-Based Learning*, *2*(3), 87–110.

Clark, D. B., Tanner-Smith, E., & Killingsworth, S. (2016) Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research*, *86*(1), 79-122.

Connolly, T., Boyle, E., MacArthur, E., Hainey, T. & Boyle, J. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, *59*(2), 661–686.

Greenspan, B., & Whitson, J. (2013). Whether to Play or Preserve the Past? Creating The Forgotten Worker Quest. In G. Yannakakis, E. Aarseth, K. Jorgenson, & J. Lester (Eds.), *Proceedings of the 8th International Conference on Foundations of Digital Games* (pp. 298–305). New York, NY: ACM.

Habgood, M.P.J., & Ainsworth, Shaaron, E. (2011). Motivating children to learn effectively: Exploring the value of intrinsic integration in educational games. *Journal of the Learning Sciences*, *20*(2), 169–206.

Horn, M., Leong, Z., Block, F., Diamond, J., Evans, E. M., Phillips, B., & Shen, C. (2012). Of BATs and APEs: An interactive tabletop game for natural history museums. In J. Konstan, E. Chi, & K. Hook (Eds.), *Proceedings of the 2012 ACM SIGCHI Conference on Human Factors in Computing Systems* (pp. 2059–2068). New York, NY: ACM.

Horn, M.S., Solovey, E.T., Crouser, R.J., & Jacob, R.J.K. (2009). Comparing the use of tangible and graphical programming languages for informal science education. In D. Olsen, R. Arthur, K. Hinckley, M. Morris, S. Hudson, & S. Greenberg (Eds.), *Proceedings of the 2009 ACM SIGCHI Conference on Human Factors in Computing Systems* (pp. 975–984). New York, NY: ACM.

Isbister, K., Flanagan, M. & Hash, C. (2010). Designing games for learning: Insights from conversations with designers. In E. Mynatt, G. Fitzpatrick, S. Hudson, K. Edwards, & T. Rodden (Eds.), *Proceedings of the 2010 ACM SIGCHI Conference on Human Factors in Computing Systems* (pp. 2041–2044). New York, NY: ACM.

Jenson, J., & De Castell, S. (2010). Gender, simulation, and gaming: Research review and redirections. *Simulation & Gaming*, 41(1), 51-71.

Johnson, W.L. (2010). Serious Use of a Serious Game for Language Learning. *International Journal of Artificial Intelligence in Education*, *20*(2), 175–195.

Kelley, J. F. (1983). An empirical methodology for writing user-friendly natural language computer applications. In *Proceedings of the 1983 ACM SIGCHI Conference on Human Factors in Computing Systems* (pp. 193–196). New York, NY: ACM.

Lane, H.C., Cahill, C., Foutz, S., Auerbach, D., Noren, D., Lussenhop, C., & Swartout, W. (2013). The Effects of a Pedagogical Agent for Informal Science Education on Learner Behaviors and Self-efficacy. In H.C. Lane, K. Yacef, J. Mostow, P. Pavlik (Eds.), *Proceedings of the 16th International Conference on Artificial Intelligence in Education* (pp. 309–318). Heidelberg, Germany: Springer.

Linehan, C., Kirman, B., Lawson, S. & Chan, G. (2011). Practical, appropriate, empirically-validated guidelines for designing educational games. In D. Tan, G. Fitzpatrick, C. Gutwin, B. Begole, W. Kellogg (Eds.), *Proceedings of the 2011 ACM SIGCHI Conference on Human Factors in Computing Systems* (pp. 1979–1988). New York, NY: ACM.

Mitgutsch, K., & Alvarado, N. (2012). Purposeful by design? A serious game design assessment framework. In M. El-Nasr, M. Consalvo, & S. Feiner (Eds.), *Proceedings of the 7th International Conference on Foundations of Digital Games*, (pp. 121–128). New York, NY: ACM.

National Research Council. 2009. *Learning science in informal environments: People, places, and pursuits*. Washington DC: The National Academies Press.

Rowe, J., McQuiggan, S., Robison, J., & Lester, J. (2009). Off-Task Behavior in Narrative-Centered Learning Environments. In V. Dimitrova, R. Mizoguchi, B. du Boulay, & A. Graesser (Eds.), *Proceedings of the 14th International Conference on Artificial Intelligence in Education* (pp. 99–106). Amsterdam, the Netherlands: IOS Press.

Schulman, D., Sharma, M., & Bickmore, T. (2008). The identification of users by relational agents. In L. Padgham & D. Parkes (Eds.), *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multi Agent Systems*, (pp. 105–111). Richland, SC: IFAAMAS.

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. (2013). A Meta-Analysis of the Cognitive and Motivational Effects of Serious Games. *Journal of Educational Psychology*, *105*(2), 249–265.