Games of Bones: Design Decisions and Early Feedback from a Prototype

Kenneth Angielczyk, Audrey Aronowsky, Beth Sanzenbacher, Johanna Thompson, Krystal Villanosa,
The Field Museum of Natural History, 1400 S. Lake Shore Dr., Chicago, IL 60605
kangielczyk@fieldmuseum.org, aaronowsky@fieldmuseum.org, bsanzenbacher@fieldmuseum.org, jthompson@fieldmuseum.org, kvillanosa@fieldmueum.org

Abstract: In order to enliven "life through time" exhibits that are standard in most natural history museums, an interdepartmental team from the Field Museum of Natural History is creating a web-based game called *Game of Bones* (GoB). GoB is implemented in Unity 3D for maximum flexibility and is intended for use both within museum exhibits and remotely. When completed, GoB will educate museum visitors and online learners about basic anatomy and evolution through ten game levels that map to seminal moments in Earth's history that are represented in almost all "life through time" exhibits. Gameplay will replicate the activities of paleontologists with players digging up fossils, re-assembling ancient animals and plants, using museum collections to test basic hypotheses about the organisms' ecologies, and making virtual museum exhibits. Here, we discuss design decisions and focus group feedback from the initial prototype of the game.

Introduction: Enlivening "Life through Time" Exhibits

Many natural history museums feature "life through time" exhibits. In these exhibits, a visitor walks past dioramas that reconstruct communities at particular points in Earth's history, presenting a broad overview of evolution and the history of life. A typical visitor to such exhibits might focus on the more charismatic reconstructions (Tubutis, 2005), like dinosaurs of the Mesozoic Era. Indeed, dinosaurs attract so much visitor attention that many museums devote the greatest amount of square footage to them in "life through time" exhibits and often separate them out into their own sections or complementary exhibits (e.g., the American Museum of Natural History or the Denver Museum of Nature and Science). The Field Museum of Natural History's (FMNH) "life through time" exhibit, Evolving Planet, opened in 2006 after a major re-design, and features over one thousand fossil specimens, dioramas, videos, and reconstructions of fossil species. The exhibit covers life from its origins to the last ice age and features a large space devoted to dinosaurs. Visitor time spent in these dinosaur sections is often at the expense of other equally important but lesser known periods in Earth's history.

An interdepartmental FMNH team is creating a web-based game that can complement and augment the visitor experience in a "life through time" exhibit such as Evolving Planet. This game will enliven the exhibit and bring attention to all periods of Earth's history, not just ones with charismatic animals such as dinosaurs. A primary goal for this prototype game is to expand players' experiences by giving them insight into the research performed by museum scientists and emphasize the key role that natural history collections have in scientific inquiry. Once completed, the 10-level game will focus on seminal moments in Earth's history. Given that "life through time" exhibits showcase similar seminal content, a major design goal is broad-scale usability of the game by museums and science centers around the country, and possibly the globe.

Game of Bones

To test the concept and design for a "life through time" game, FMNH produced a prototype called *Game of Bones* (GoB). GoB is aimed at middle- and upper-elementary school aged youth and families with the goals of allowing players to experience what it is like to be a paleontologist by replicating the activities of scientists that work in museums through engagement with museum research, collections, and exhibits. GoB has been prototyped as a single-player experience to focus the narrative and test different designs and basic mechanics. A collaborative or multi-player mode will be tested in future design iterations because collaborative gameplay would more closely mimic real scientific experience and promote greater learning. In the single-player prototype, GoB players unearth, prepare, and study important fossils in Earth's history. Players also gain insight into the nature of scientific inquiry through topics such as anatomy, functional morphology, and evolution. Designed in Unity 3D, GoB will be made available at no cost through the Museum's website and will be adapted for tablet devices to enable gameplay from within museum exhibits.

GoB is based on current and past paleontological research conducted by FMNH scientists, with an emphasis on scientific accuracy and the realities of field- and museum-based research. Designing games around real scientific research with a high degree of accuracy is critical for several reasons: 1) it simulates real-life scientific observations and experiences for players; 2) it allows players to mimic scientific processes in order to generate solutions to real-world questions; 3) it provides real-life scientific discovery moments and opportunities for higher-level engagement (Aronowsky et al, 2011). In the finished game, youth will use gameplay to learn how paleontologists work, investigate fossil specimens in museum exhibits and research collections, and gain insight into the nature of scientific inquiry and topics such as geography, evolution, anatomy, and functional morphology. We also seek to engender a positive attitude towards science in youth, a goal that is realistic given past success with paleontology-based digital learning programs such as I Dig Science (King et al, 2011; Steinkuehler and Alagoz, 2010). The GoB game design and mechanics are targeted towards youth aged 9–13 (approximately 4th–8th grades), but we anticipate that GoB will also appeal to family audiences. We envision the final game as having ten levels with varying degrees of difficulty. Each level will correspond to a different period in geologic time and will focus on an FMNH specimen and the research of an FMNH curator or collections manager. The prototype focuses on *Edaphosaurus* (Figure 1), a fossil mammal-relative from the Early Permian Period of Earth history (approximately 299 to 270 million years ago) and the research of Assistant Curator of Paleomammalogy Dr. Kenneth Angielczyk.

The Museum's departments of Education, Geology, and Biodiversity Synthesis Center worked with two recent graduates from Columbia College Chicago (a game designer and a graphic artist) to create the game prototype. The prototype is intended to exemplify the vision of the entire game by showing one fully playable stage that presents examples of game mechanics, graphics, science content, and learning goals. The prototype incorporates aspects of the virtual dig, research, and museum exhibit activities inspired by the I Dig Science summer program (see below), with the same attention paid to the scientific accuracy and realism of paleontological fieldwork and museum research.

Game of Bones

The inspiration for GoB's design is the successful I Dig Science summer camp, an out-of-school time program for a small group of high school-aged youth that uses a suite of digital technologies in combination with real-world resources at FMNH. I Dig Science was created by FMNH and Global Kids in 2008 and provides opportunities for teens from disparate locations to use a participatory 3D virtual environment to communicate with scientists and conduct activities that mimic those of the scientists—including hunting for fossils, collecting data, testing hypotheses, and discussing and interpreting their discoveries. The GoB prototype is an attempt to engage larger groups of younger learners and families in a web-based game that captures some of the core science concepts and learning goals of I Dig Science.

In the prototype, players begin gameplay in a virtual museum office, based on a real space in the FMNH Geology Department. The office serves as the hub for all activities within the game and also provides players with a virtual behind-the-scenes look at the Museum. Our choice to use the office as a hub for the game was based on 1) the fact that it adds to the game's immersive environment, making players feel that they are scientists working in a museum; and 2) the contents of a typical curator's office (e.g., maps, references, specimen cabinets, workspace for studying fossils) consist of many items that lend themselves naturally to certain aspects of gameplay and that players will actively use in the game. After clicking the map (Figure 1), players are presented with a series of virtual fossil excavation sites around the globe to which they can travel. These dig sites are based on the real locations where FMNH researchers have ongoing projects or have worked in the past, emphasizing the diversity of work performed by Museum scientists. In the prototype, only one dig site can be selected. However, in the completed version of the game, a subset of localities will be available to the player at the start of the game, with additional localities becoming unlocked as players gain experience as a paleontologist. For the prototype, we chose to implement an example focusing on the mammal-relative *Edaphosaurus* (Figure 1), which is commonly found in Lower Permian rocks in Texas and Oklahoma (e.g., Reisz, 1986; Berman et al., 1997). Our decision to use *Edaphosaurus* reflects a large number of factors: 1) its distinctive and engaging morphology (e.g., it has a large sail on its back), 2) its evolutionary importance as one of the first terrestrial vertebrate herbivores (e.g., Reisz and Sues, 2000; Reisz, 2006), 3) the large number of *Edaphosaurus* specimens in FMNH collections, 4) a skeleton of *Edaphosaurus* is displayed in Evolving Planet and in most "life through time” exhibits, and 5) Dr. Angielczyk’s research focuses on the paleobiology of ancient mammal-relatives (non-mammalian synapsids) such as *Edaphosaurus*.

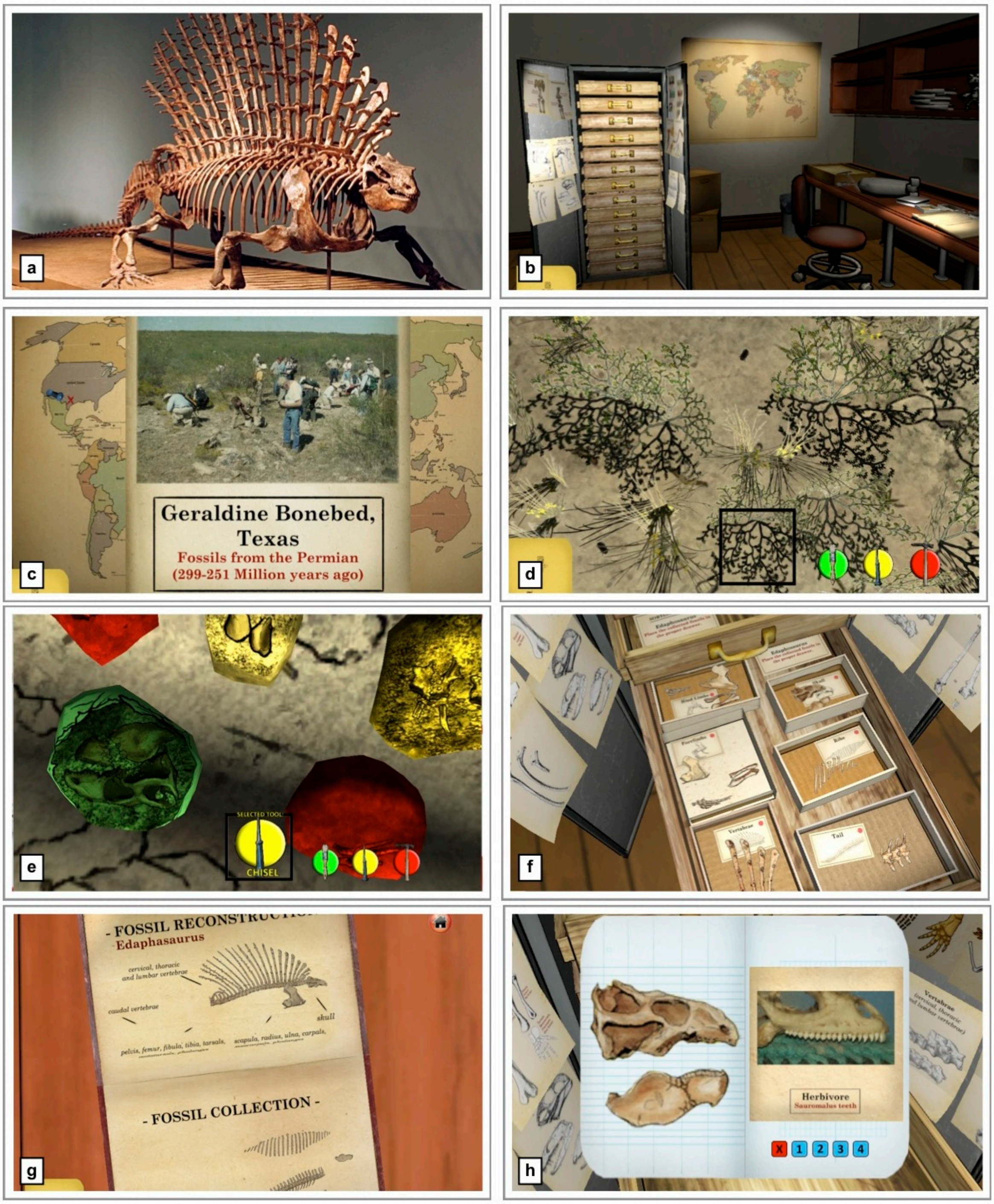

Figure 1: **Images from Game of Bones (from top to bottom, left to right) a. *Edaphasaurus* specimen from Field Museum, b. Office, c. Map, d. Dig site, e. Digging, f. Sort, g. Assemble, h. Research**

Once a locality is selected, players hunt for fossils at their site, using the same tools that paleontologists use to excavate fossils in that area. For example, when digging in Texas, players use rock hammers, chisels and brushes, but if digging in the Antarctic, players will use jackhammers and saws. This portion of the game allows players to experience how paleontologists discover and excavate fossils, and understand the types of tools needed for fieldwork. The fossil excavation game mechanic should be compelling to players because it can conjure the sense of excitement and adventure learners often associate with paleontological fieldwork in exotic locations. The inclusion of the virtual fossil dig in GoB also provides an accessible gateway to the game that may inspire players to further engage with it, even after they discovered their fossils.

After players have successfully excavated their fossils, they return to the office to sort and assemble their discovery. To maintain the exploratory feel of the game and prevent it from becoming overly linear, players can choose which task they want to accomplish first. When sorting, they use comparisons with real references from the scientific literature to help them sort their fossil pieces into the correct anatomical bins (e.g. skull, shoulders, pelvis). During assembly, players use their fossil pieces to reconstruct their ancient animal. Our goal for these mechanics is in part to enrich the simulation of being a scientist, since organizing, preparing, and identifying fossils are important steps in the follow-up work after an expedition. The sorting process also emphasizes how scientists use previous work in the literature to inform their research, and provides a seamless way for players to begin thinking about the anatomy of their fossil animal and how it may be similar to or different from the anatomy of other fossil and extant animals. In particular, the comparisons with literature sources emphasize our homology learning goal (i.e., homologous structures are equivalent because of their common evolutionary origin even if they appear somewhat different due to functional overprinting). The assembly mechanic gives players the opportunity to apply what they may already know about anatomy, or have learned during the sort process, with the visual reward of seeing their animal come together through their efforts.

Once they have sorted and assembled their fossil animal, players are challenged to develop and test hypotheses about its biology, ecology, and functional morphology. To provide insight into the key role that museum specimens play in scientific research, one of the main mechanics of this stage of the game consists of players using the “research collection”, where they find photographs of relevant parts (e.g., skulls, jaws, limb bones) of both extinct and extant animals. By making observations of important features (e.g., tooth and skull shape in regards to dietary preference), and synthesizing these data with additional information provided during the game and knowledge that they already have, players can make inferences about the lifestyle of their fossil animal. In the prototype, gameplay in this stage focuses on answering the question “What did your animal eat?” This allows players to experience how paleontologists reconstruct the biology and ecology of extinct animals and to also gain insight into topics such as the form-function relationship that exists in many organisms (e.g., shearing teeth are needed to slice meat). Although our focus group testing (see below) showed that the game mechanics of this stage are fundamentally sound, time and funding constraints limited our implementation to only including photographs of comparative specimens. As our development of the full game proceeds, we hope to replace the photographs with fully-rendered objects comparable to the fossil specimens that players excavate and study, which will help to make this stage more engaging.

The game concludes with an often overlooked aspect of the scientific process: presenting one’s results and conclusions to peers and the general public. This mechanic takes the form of players building a virtual museum exhibit about their fossil animal. The museum exhibit format was chosen because it combines a way for players to integrate all aspects of their previous gameplay (including their fully reconstructed fossil and their conclusions about its biology and ecology) in a format that is interesting, familiar, and that ties their experiences back to a museum context.

Youth Focus Group Feedback and Surveys

The preliminary learning goals of GoB are for players to understand paleontological fieldwork and research, what museum collections are and how they are used, basic anatomy and the principle of homology, the functional morphology of living things, and the nature of scientific inquiry. Table 1 outlines eight preliminary learning and attitudinal goals for GoB and how they map to gameplay in the prototype. These goals may be modified or expanded in future iterations of the game.

Learning Goal	Game Mechanic
1. Understand the distribution of fossils in time and space	Using Map
2. Understand the realities of paleontological field work – types and conditions of fossils	Digging Fossils
3. Know some of the tools & technology used for paleontological dig	Digging Fossils
4. Understand basic anatomy and the principle of homology	Sorting, Assembling & Researching Fossils
5. Apply logic/comparative methods to solve a problem	Sorting & Assembling Fossils
6. Understand what museum collections are and how they are used	Researching Fossils
7. Understand the function of living things (functional morphology)	Researching Fossils
8. Have a positive effect towards science, feel a empowered in science	Overall Game Play

***Table 1*: Game of Bones Learning Goals**

An early version of the prototype was tested with a youth focus group of twenty-three 6th grade students in November 2011. We administered pre- and post-play surveys to better understand what youth players were gleaning from gameplay. Fourteen of twenty-three (60%) players had a positive change in their understanding and description of the realities of paleontology fieldwork (learning goal #2). All but one participant demonstrated an understanding of the tools used by paleontologists post-gameplay (learning goal #3). All but two participants showed positive change in knowledge about vertebrate anatomy (learning goal #4), and these two participants had no change between the surveys as they demonstrated a sound comprehension of anatomy in the pre-play survey. All players were able to apply logic in problem solving (learning goal #5), although 83% (19/23) demonstrated this skill in the pre-play survey. All but one player were able to articulate why museums have and need research collections in the post-play survey (learning goal #6). Seventy percent of players showed positive change in their ability to evaluate anatomical function post-play (learning goal #7). There were no differences in attitudes towards science because both pre- and post-play survey responses were uniformly positive towards science with all youth able to articulate specific ways in which science benefits the world (learning goal #8). Given that the prototype focuses on a single geological time period, we did not evaluate students' understanding of geological time and the distribution of fossils in time and space (learning goal #1).

From these preliminary data, we see areas where the current design of GoB can have a positive impact on learning. Gameplay may help players to understand 1) the realities of paleontology; 2) the tools and methods used by paleontologists; 3) the importance of museum collections for science; 4) basic anatomy; and 5) the ability to evaluate anatomical function. Based on preliminary data, GoB might not have a positive impact on a player's attitude towards science or ability to apply logic to solve a problem because interest-driven players may already have achieved these two goals.

For the youth focus group, we solicited feedback about gameplay and mechanics (what worked and what did not) and if and where players envisioned themselves playing the completed ten-level game. Critical feedback from youth players indicated they wanted to play more and have more levels (the most common comment was "make the game longer"). Most youth players indicated that they would enjoy playing a full version of GoB from home, which is interesting considering that these same players most frequently played sports games such as *Madden 12*, shooter games such as *Call of Duty: Black Ops*, and physical games such as *Just Dance 2* at home. Very few youth players indicated they would play GoB in a museum setting. More focus groups and a different testing design would be necessary to tease apart 1) if youth do not currently associate natural history museums with digital gameplay; 2) if youth players cannot envision playing a digital game from within a museum because they currently lack the technology to do so; or 3) if youth would not play a digital game from within a natural history museum even if provided with the necessary technology.

Adult Focus Group

Thirteen museum staff members from the Geology, Education, and Exhibits departments participated in an adult focus group in December 2011. For the adult focus group, we were interested in how well the gameplay mimicked the experience of a paleontologist, how to mesh gameplay and an accurate depiction of the scientific process, the look and feel of the game, and suggestions for expansion and improvement. These adult testers said the game concept was both engaging and educational. Players stated that the game activities and mechanics accurately mimicked scientific activities and presented the underlying concepts and facts in an appropriate manner for the intended audience. Most adults who participated were not avid gamers and suggested that an introductory tutorial would help to orient them to the game storyline and mechanics. These testers did not appear to recognize standard game design elements such as guiding a player through a space using lighting, and had a harder time navigating the game than did youth players. Although our conclusions are limited by a very small sample size and biased due to the source of the sample (all museum staff), adult and youth players favored different aspects of the game prototype. Adult players seemed to favor the graphics and accuracy of GoB, naming the Office, Notebook, and Reconstruction as their favorite parts of the GoB prototype. In contrast, youth players preferred the activities, naming Fossil Assembly and Sorting as the most engaging parts of the game.

Future Development

As iterative development continues on GoB, the FMNH team will take a two-pronged approach: 1) refine single-player web-based gameplay and 2) explore and prototype collaborative and in-museum gameplay. A single-player non-museum game was the preferred product of the youth focus group, so it is important to continue developing this design. However, testing the feasibility of a multi-player non-museum game will also be a priority for future development. Given our ultimate goal of broad-scale usability at museums nationally, in-museum gameplay will be prototyped as well. FMNH expects in-museum development to include a new and different suite of focus groups, site visits to institutions around the country, and discussions with exhibits and paleontology staff.

Conclusion

Science content and the scientific process are often inaccessible for youth. Youth often perceive science as a collection of obscure facts that are unrelated to their daily lives, and do not realize that science is a dynamic activity. At its most basic, science entails asking questions and making observations, two activities that people practice regularly, if unknowingly, in their daily and digital lives. When complete, the scientifically accurate GoB may provide a way for players to apply or improve these skills in context, making the scientific process accessible and familiar. Gameplay may increase content knowledge, heighten interest in science, and engender positive attitudes toward science among players. Paleontology has a strong feeling of adventure and discovery that may help to draw in youthful players who might otherwise be reluctant to engage with science content. By making science accessible and increasing their content knowledge, the scientific experience provided by GoB has the potential to provide youth players with a better understanding of paleontology and museum research.

References

Aronowsky, A., Sanzenbacher, B., Thompson, J., and Villanosa, K. (2011). Worked example: How scientific accuracy in game design stimulates scientific inquiry. *The International Journal of Learning and Media*, 3(1). doi: 10.1162/IJLM_a_00064.

Berman, D.S., Sumida, S.S., and Lombard, R.E. (1997). Biogeography of primitive amniotes. In S.S. Sumida and K.L.M. Martin (eds.), *Amniote origins: Completing the transition to land*, pp. 85-139. San Diego: Academic Press, San Diego.

King, E., Alagoz, E., Martin, C., Chu, S., Zhang, B., Oh, Y., & Steinkuehler, C. (2011). Soft modding in two out-of-school virtual worlds-based programs. *The International Journal of Learning and Media*, 2(4). doi: 10.1162/ijlm_a_00054.

Reisz, R.R. (1986). Pelycosauria. *Handbuch der paläoherpetologie*. Teil 17A. Stuttgart: Gustav Fischer Verlag.

Reisz, R.R. (2006). Origin of dental occlusion in tetrapods: Signal for terrestrial vertebrate evolution? *Journal of Experimental Zoology (Molecular Development and Evolution)*, 306B: 261-277.

Resiz, R.R. and Sues, H.D. (2000). Herbivory in late Paleozoic and Triassic terrestrial vertebrates. In H.-D. Sues (ed.), *Evolution of herbivory in terrestrial vertebrates: Perspectives from the Fossil record*, pp. 9-41. Cambridge: Cambridge University Press.

Steinkuehler, C. & Alagoz, E. 2010. Out-of-school virtual worlds based programs: A cross-case analysis. In K. Gomez, L. Lyons, J. Radinsky (Eds.) *Learning in the Disciplines: Proceedings of the 9th International Conference of the Learning Sciences.* (ICLS 2010) Volume 2, Short Papers, Symposia, and Selected Abstracts (pp. 304-305). International Society of the Learning Sciences: Chicago IL.

Tubutis, T., 2005. Revitalizing “Life over time”: A new look for a very old topic. *In The Field* 76(2): 18.

Commercial Video Games as Preparation for Future Learning

Dylan Arena, Stanford University, darena@alumni.stanford.edu

Abstract: To examine the learning benefits on traditional school content of recreational play of commercial games, 102 community-college participants were randomly assigned to play *Call of Duty 2*, *Civilization IV*, or no game at home for at least 15 hours over 5 weeks. All participants then took a short multiple-choice test about World War II history; heard a 20-minute lecture about World War II; and then took another multiple-choice test about World War II history. Results (using an intention-to-treat analysis) showed no differences on the pre-lecture test but a positive effect (Cohen's d = .27) of gameplay on the post-lecture test, suggesting that recreational gameplay had prepared participants to learn from the lecture. These findings suggest a new role for games in learning contexts, in which the games—instead of carrying the educational load alone—provide compelling experiences that are coupled with the powerful explanatory structures of a formal curriculum.

Introduction

In the field of digital game-based learning there has been much focus on the development of educational games—games designed specifically to teach. Attention has also been paid to commercial, off-the-shelf (COTS) games, but that work has focused mostly on the motivation, engagement, and community participation engendered by such games. To complement such work, in this paper I hope to demonstrate that recreational COTS gameplay can, under the right circumstances, lead to learning gains on even the traditional, fact-based curricular instruments criticized by many in the educational research community but used widely in schools today. By showing this learning benefit of COTS games, I hope to strengthen the argument for digital game-based learning more broadly.

It is not surprising that there has been relatively little research demonstrating the benefits that might accrue on traditional fact-based tests from simply playing COTS games in one's leisure time, because there is no reason to expect that games created purely to entertain would produce learning benefits on school content that are measurable by traditional assessments. A novel assessment framework called Preparation for Future Learning (PFL), however, is designed specifically to measure immature forms of knowledge that traditional assessments miss. It does this by incorporating learning resources into the assessment process to determine what test takers' immature knowledge has prepared them to learn (Bransford & Schwartz, 1999).

In this study, I used the PFL framework to investigate whether being randomly assigned to receive and play one of two COTS games (*Call of Duty 2*, n = 34, or *Civilization IV*, n = 35) at home over the course of five weeks would prepare community-college students to learn from a lecture about World War II compared to a control condition that received no game (n = 33), as measured by performance on pre- and post-lecture multiple-choice tests. In a more qualitative vein, I also examined whether participants' gameplay experiences would influence their responses to open-ended questions about scenarios from World War II that were not mentioned in the lecture.

Methods Design

The study was an experimental field trial with non-random selection of participants from a convenience sample but random assignment of participants to three conditions. Two of these conditions' participants were given one of two commercial video games, *Call of Duty 2* (*CoD2*) or *Civilization IV* (*Civ4*), which they were assigned to play at home for at least 15 hours over a period of five weeks; the third condition's participants were given no game and assigned no gameplay. After this five-week period, participants from all three conditions came to a room on their community-college campus, took a 16-item pretest about World War II history, watched a 20-minute video of a narrated-slideshow lecture about World War II history, and then took a 36-item posttest and a brief survey. After completion of the session, participants were given research credit, and participants in the condition that had not already received a game were given one and asked to play it for at least 15 hours over a five-week period. (Although the post-instructional gameplay of these control-condition participants was not of theoretical interest for this study, I did not want control participants to participate less fully because of being assigned to a condition that received neither a game nor the opportunity for a gift card, so I explained in the initial information session that participants in all three conditions would receive games to play for compensation and that the only condition difference would be the timing of that gameplay.) All participants who completed sufficient gameplay (as determined by examination of save-game files) were e-mailed digital $75 Amazon.com gift cards. Figure 1 details this design.

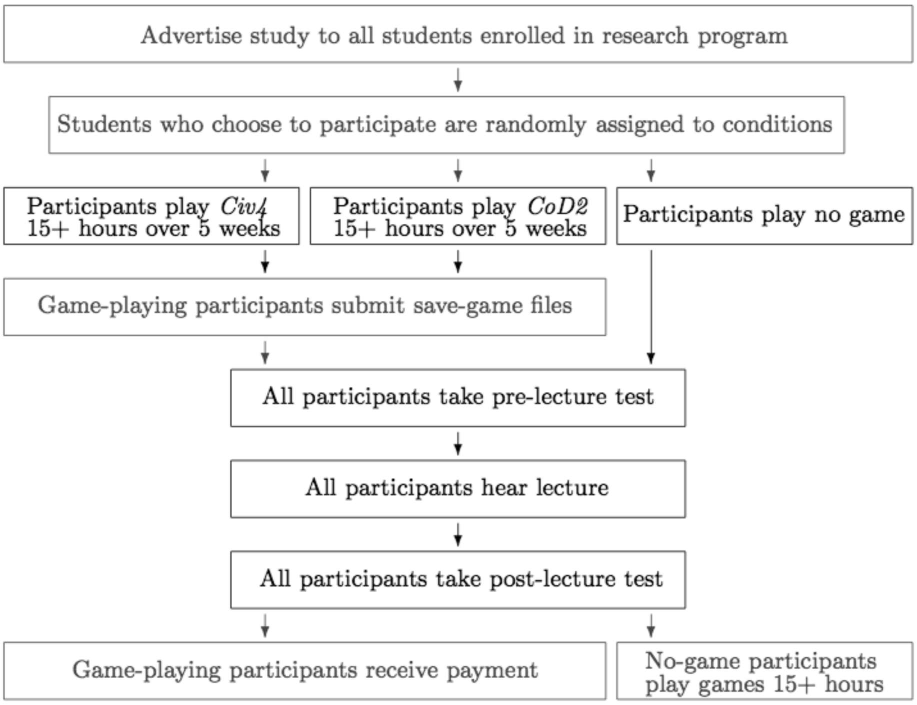

***Figure 1*: Experimental design and procedure.**

Participants

I recruited 119 participants and obtained usable data from 102(1), all of whom were students at a local community college who were enrolled in introductory social-science classes that required research participation for course credit. The final participant sample exhibited high diversity in ethnicity, nationality, socioeconomic status, and prior digital gameplay. Participants ranged in age from 16 to 42 years, with a median age of 20, and 66 of the 102 participants were female.

Materials

Each participant was given a new, shrink-wrapped copy of either *CoD2* or *Civ4*, installable on the participant's personal computer (both Windows and Macintosh versions were available).

The lecture took the form of my narration of a 24-slide presentation that discussed the events of World War II while focusing on the two themes of *Nations* and *Battles*, which I hypothesized to be more relevant to the players of *Civ4* and *CoD2*, respectively. Although the lecture was written so that players of the two games would retain information from different parts of it, I never mentioned either game in the lecture.

The multiple-choice-tests, which were intended to look for benefits of recreational gameplay even on the most traditional, "schoolish" measures, comprised items from the National Assessment of Educational Progress (NAEP: National Center for Education Statistics, 2011), the California Standards Test (CST: California Department of Education, 2009a and 2009b), and a purpose-built test from a World-War-II study guide produced by the company SparkNotes (SparkNotes Editors, 2005), which also served as a primary source for the lecture.

I also administered a pre-experimental questionnaire and an exit survey to (a) screen out anyone who had played either game before, (b) collect demographic data, and (c) solicit feedback about game enjoyment and learning behaviors.

As another way to examine effects of gameplay on learning, I asked participants open-ended questions about two scenarios presented in the exit survey after the post-lecture test. The scenarios had not been mentioned in the lecture but were intended to build upon the lecture's two themes of *Nations* and *Battles*, respectively. In the *Nations* scenario, British ships fired on their French allies in 1940 off the coast of Algeria. In the *Battles* scenario, American soldiers scaled a cliff under heavy German fire to disable an artillery battery. After presenting each scenario, I asked participants what they would want to learn to better understand the scenario.

Results

My primary analysis protocol was as follows. I first identified a set of variables that might predict performance on the pre- and post-lecture tests, as shown in Table 1.

	Control	*CoD2*	*Civ4*
Age in years			
M (SD)	22.76 (6.77)	22.65 (6.84)	22.60 (5.94)
English proficiency (reading/speaking composite)			
M (SD)	3.17 (1.23)	3.53 (1.71)	3.31 (1.56)
Game enjoyment			
1, 2, 3, 4, 5 (on Likert scale)	0, 0, 33, 0, 0	3, 5, 7, 11, 8	2, 6, 8, 15, 4
M (SD) [if treated as continuous]	3 (0)	3.47 (1.26)	3.37 (1.09)
Gender			
Females, Males	21, 12	21, 13	24, 11
Prior digital gameplay			
Never, 1-2 times, 3-6 times, > 6 times	5, 4, 7, 17	6, 7, 3, 18	3, 2, 4, 26
Prior social-studies interest			
1, 2, 3, 4, 5 (on Likert scale)	2, 1, 8, 13, 8	2, 3, 8, 15, 6	2, 4, 8, 11, 10
M (SD) [if treated as continuous]	3.73 (1.07)	3.59 (1.08)	3.66 (1.19)
Quarter			
Winter, Spring, Autumn	17, 7, 9	17, 7, 10	13, 13, 9

***Table 1*: Candidate predictor variables.**

Next, I constructed Analysis of Covariance (ANCOVA) models with the pre- and post-lecture-test scores as dependent variables and the variables listed in Table 1 as predictors, along with a gameplay-condition factor (because my primary interest for these analyses was the effect of gameplay rather than of each specific game, this two-level factor contrasts *CoD2* and *Civ4* participants with Control participants). For the post-lecture test, I also included pre-lecture-test score. My first step for each analysis was to test a model containing main effects of all predictors against a model that also contained all one-way interactions with the gameplay-condition factor, but in neither case did the marginal explanatory power of the interaction model reach statistical significance. Therefore, neither of the models discussed contains any interaction terms. Using the saturated main-effects model as a starting point, I then performed an all-possible-subsets model selection (2) to find the model with the highest R^2_{adj}, with the constraint that all predictors in models under consideration be at least marginally significant ($p < .1$). I will present these "parsimonious" models below.

Pre-lecture test

As shown in Figure 2, mean scores on the pre-lecture test were low for all three conditions.

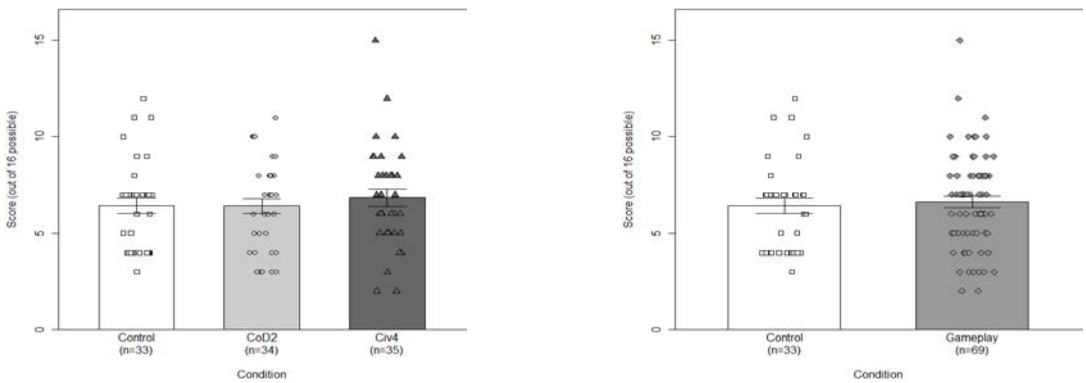

***Figure 2*: Pre-lecture-test scores by condition and gameplay.**

The test also had low reliability for this sample (Cronbach's α = .43), most likely because many participants were guessing on many of their response, thus introducing random noise to the measurement. The sole predictor chosen by the all-possible-subsets selection procedure for the parsimonious ANCOVA model for pre-lecture-test scores was English proficiency, as shown in Table 2. Notably, gameplay condition was not predictive.

	df	$SS_{TypeIII}$	***F***	η^2	***p***
English proficiency	1	70.86	13.32	.12	.00042***
Residuals	100	532.16			
R^2_{adj} = .11, $F(1, 100) = 13.32$, p = .00042					

***Table 2*: Pre-lecture test ANCOVA.**

Post-lecture test

As shown in Figure 3, scores on the post-lecture test were much higher than on the pre-lecture test, indicating that participants in all conditions learned from the lecture. The test's reliability for this sample was also much higher than was the pre-lecture test's (Cronbach's α = .86). Mean scores in the gameplay conditions were slightly higher than in the control condition, suggesting a small benefit of gameplay (Cohen's d = .27).

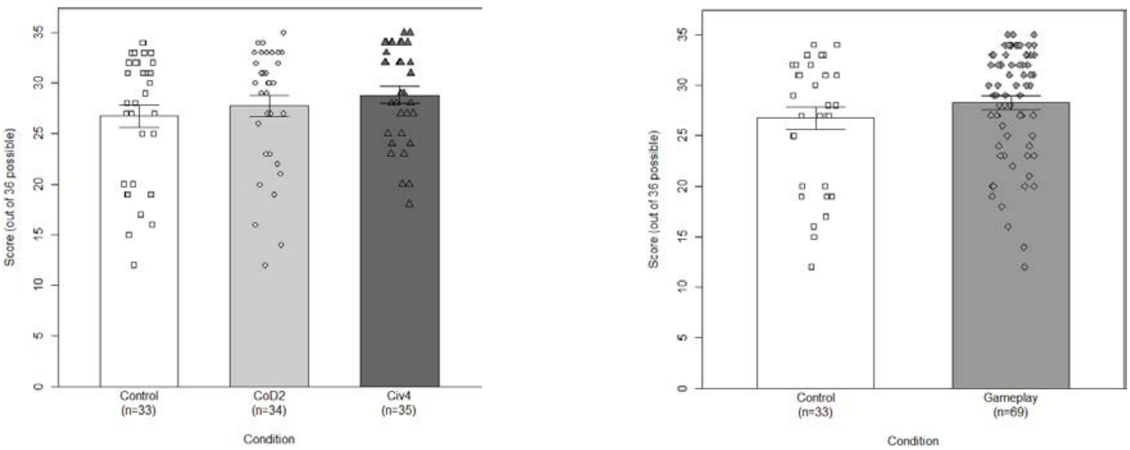

***Figure 3*: Post-lecture-test scores by condition and gameplay.**

The all-possible-subsets selection procedure for the ANCOVA model with post-lecture-test score as dependent variable produced a model with seven predictors, as shown in Table 3. Notably, gameplay condition (denoted by the "Received a game" variable) was selected in this model, albeit with a small effect size (η^2 = .026). Age, English proficiency, game enjoyment, pre-lecture-test score, and prior social-studies interest were also positively associated with post-lecture-test score. (Quarter of participation was marginally significant, reflecting cohort or seasonality effects that are not of theoretical interest but that contribute construct-irrelevant variance.)

	df	$SS_{TypeIII}$	***F***	η^2	***p***
Age	1	123.55	5.63	.036	.020*
English proficiency	1	70.86	13.32	.12	.00042***
Game enjoyment	4	239.07	2.72	.069	.035*
Pre-lecture-test score	1	145.08	6.60	.042	.012*
Prior social-studies interest	4	334.12	3.80	.096	.035*
Quarter	2	125.40	2.85	.036	.063.
Received a game	1	88.97	4.05	.026	.047*
Residuals	87	1910.95			
R^2_{adj} = .36, $F(14, 87) = 5.06$, $p < .0001$					

***Table 3*: Post-lecture test ANCOVA.**

Open-ended questions
Because I had intended the *Nations* theme to resonate more with *Civ4* players and the *Battles* theme to resonate more with *CoD2* participants, I predicted that the *Nations* scenario would elicit more responses focused on *Nations* themes from *Civ4* players, and likewise the *Battles* scenario would elicit more *Battles*-focused responses from *CoD2* participants. Figure 4 shows results by condition for the two scenarios. In these graphs, correctness was determined by whether the questions participants asked about the scenario reflected a focus on the appropriate theme for that scenario: e.g., for the *Nations* scenario, a question about whether the French ships being fired upon were controlled by Germany would be scored as reflecting a *Nations* focus, whereas a question about whether the French commanders had insulted the British commanders would not. The error bars in each graph represent the 68% confidence intervals for the proportions (corresponding to roughly +/- 1 SD, assuming normality).

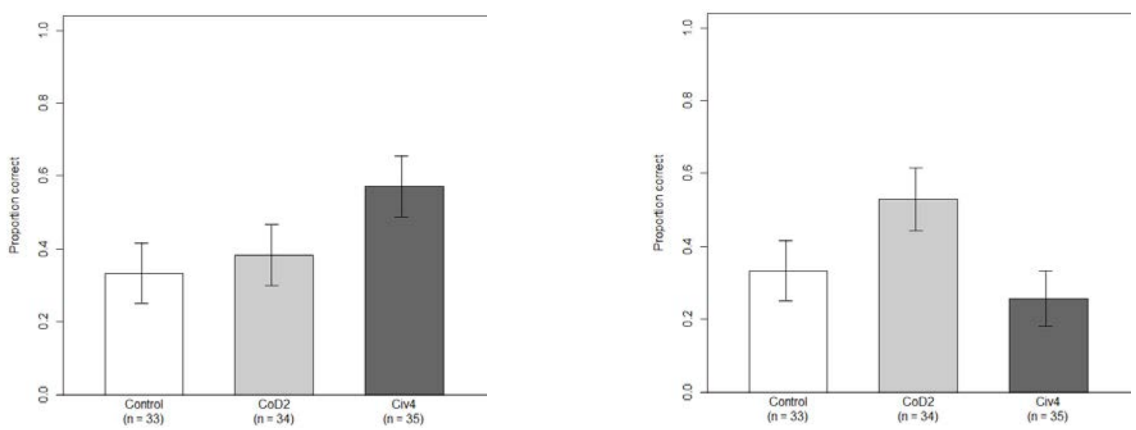

***Figure 4*: Open-ended responses for *Nations* (left) and *Battles* (right) scenarios.**

Fisher's exact test of proportions for the *Nations* scenario was marginally significant, p = .058; for the *Battles* scenario, the test was significant, p = .030.

Discussion
Taken together, the results of this study support the claim that playing enjoyable video games at home can help both male and female students learn in school, if the formal instruction leverages the students' gameplay experiences. (The strong predictive effect of prior social-studies interest for the post-lecture test shows the importance of also leveraging students' interests.) The multiple-choice-test results showed only a small effect, but this study was intentionally conservative with respect to its design (randomized field trial using an intention-to-treat analysis, both considered "gold standard" methodologies for causal inference), its intervention (recreational gameplay of commercial games that were not designed to teach school content), its outcome (learning gains compared to control participants on traditional history content delivered via direct instruction), and its measurements (traditional multiple-choice tests whose items were taken from existing standardized tests rather than developed *ad hoc*). The open-ended-question results underscore the notion that more creative measures show stronger positive effects of recreational gameplay. They also demonstrate that different games will offer different types of experiences that prepare players preferentially for different topics of formal instruction and that these gameplay experiences can improve not only retention of facts presented by direct instruction but also students' choices about what to learn.

I propose two main conclusions from this study. From a theoretical perspective, there is a benefit to simply having demonstrated that the learning that occurs in naturalistic gameplay can be detected with the PFL framework. Showing that fruit can be plucked from this region of the digital-game-based-learning space (i.e., involving pre-instructional gameplay paired with a formal curriculum) strengthens the basic argument for digital game-based learning.

From a more pragmatic perspective, demonstrating that informal COTS gameplay can lead to learning gains in schoolish contexts suggests a specific policy prescription for educators: to consider the games their students are already playing not as competition for precious time that could be spent studying or doing homework but as rich source material for use in engaging curricula that could tie the compelling experiences found in the games with the powerful explanatory structures found in the standard curriculum. A concomitant policy prescription for commercial-game developers is that their games could contribute to efforts in digital game-based learning without having entire curricular units crammed into them—developers need only be thoughtful about how the experiences provided in their games might be tweaked this way or that to better serve as foundations upon which educators might build. This lowered bar for participation in the digital-game-based-learning space might encourage more commercial developers to lend their considerable strengths to the process of bringing classrooms into the 21st-century.

Endnotes

(1) The 16 participants who dropped out did so because either they were no longer enrolled in a relevant social-science class or because they had technical or time-management problems preventing them from completing the at-home gameplay. The 17th lost participant was removed for cheating on the post-lecture test.

(2) An all-possible-subsets selection procedure examines every model that could possibly be constructed using a set of predictor variables to find the “best” model according to some pre-defined criterion.

References

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of Research in Education* (Vol. 24, p. 61-100). Sage.

California Department of Education. (2009a). Released test questions: Grade 11 U.S. history - social science. Retrieved from: http://www.cde.ca.gov/ta/tg/sr/documents/cstrtqhssmar18.pdf on January 10, 2011.

California Department of Education. (2009b). Released test questions: World history - social science. Retrieved from: http://www.cde.ca.gov/ta/tg/sr/documents/cstrtqhssworld.pdf on January 10, 2011.

National Center for Education Statistics. (2011). NAEP questions tool. Retrieved January 10, 2011, from: http://nces.ed.gov/nationsreportcard/itmrlsx/search.aspx?subject=history.

SparkNotes Editors. (2005). Sparknote on World War II (1939-1945). Retrieved December 22, 2010, from: http://www.sparknotes.com/history/european/ww2/.

Acknowledgments

Financial support for this study was provided by a Stanford University School of Education Dissertation Support Grant and a Gerald J. Lieberman Fellowship. Theoretical support was provided by Dan Schwartz and Ed Haertel.