MARTIAN BONEYARDS

Collaborative Scientific Inquiry in *Arcadia*: An MMO gaming environment in *Blue Mars*

Jodi Asbell-Clarke, Teon Edwards, Jamie Larsen, Elizabeth Rowe, Elisabeth Sylvan

TERC

Jim Hewitt

University of Toronto
Abstract

*Martian Boneyards* is a game of scientific mystery that shows promise that social, digital games may be successful in promoting and measuring sustained scientific inquiry. Educators and game designers worked together to create this highly aesthetic and compelling game in the high-definition, massively multiplayer online environment (MMO), *Blue Mars*, striving to engage a new audience in scientific activity. A public, game-savvy audience spent an average of 30 hours in the game over the 4-month implementation period, with some players exceeding 200 hours in the game. Their time was spent primarily conducting scientific inquiry, mostly gathering data but also with about 30% of their in-game interactions in the analysis and theory-building phases of inquiry.
1. Introduction

The Educational Gaming Environments group (EdGE), a team of developers and researchers at TERC in Cambridge, MA, is looking beyond today’s schools to envision learning environments that transcend formal and informal boundaries, leveraging the learning that takes place in peoples’ everyday lives.

While vast resources are being used to develop materials for an educational system in the US that is increasingly irrelevant to students’ needs (Wagner, 2008; U.S. Dept of Labor, 2007), Internet-based free-choice environments are becoming a major source for science learning and social activity (Ito et al., 2008; Lenhart, 2010; Falk & Dierking, 2010). EdGE investigators believe that by looking at the scientific inquiry that players’ engage in within digital games, educators may learn how to harness the passion, inquisitiveness, and “blissful productivity” of gamers (McGonigal, 2011, p.53).

2. Statement of the Research Problem

Youth and adults, both male and female, are spending increasing amounts of time playing computer games. Often, these games are driven by very high-end graphical engines creating realistic and spectacular imagery (Ito et al. 2008; Lenhart, 2010). MMOs are becoming a huge new venue for socializing (Castronova, 2007; Gartner, 2008). EdGE strives to create games people choose to play in their everyday lives that are also intentionally designed to support scientific knowledge building and inquiry.

The knowledge building behaviors observed in some social virtual games have an intriguing alignment with practices of professional scientists and also with elements of
situated learning models. In games, players often work together as part of a community to solve domain-specific problems with access to informational resources and tools necessary for each problem (McGonigal, 2010). Peer-review, collaboration, sharing and analysis of data, and evidence-based reasoning are occurring in many popular role-playing games (e.g., *World of Warcraft*) (Steinkeuhler & Duncan, 2008). These gaming activities are similar to the habits of practicing scientists in professional communities who share data and observations, challenge and confirm each others claims, and work together to build theories through a well-recognized and explicit peer-review system (Dunbar, 2000).

Gamers’ activities are also suggestive of situated learning models such as the community of practice model (Lave, 1988; Lave & Wenger, 1991). The community of practice model grows from situated learning theory, which suggests that knowledge, activity, and environments are inextricably entangled (Lave, 1988). In a community of practice, people work together on domain-specific knowledge building using common habits, language, and cultural rules of engagement. Vygotsky (1978) recognized the mediating affects of community and tools on learning and described a zone of proximal development (ZPD) that is the difference between what a learner can do individually and what s/he could do with assistance from others. Interestingly, a similar tenet of many game-design models is that tasks must be just outside the current grasp of a player—doable, yet challenging—and often requiring the assistance of other players and/or tools within the game (McGonigal, 2010).

Design research projects such as Harvard University’s *River City* (Ketelhut, 2007) and Indiana University-Bloomington’s *Quest Atlantis* (Barab et al., 2005; Barab et al.,
2007; Barab et al., 2008) have used social gaming to support student communities that evolve and work together in intentional learning environments—gaming environments created for use in conjunction with formal education. In River City (Galas & Ketelhut, 2006, Ketelhut, 2007) and in Quest Atlantis (Barab et al., 2005a; Barab et al., 2005b), middle school students collaborate with classmates to solve problems critical to their community. River City has proven successful in engaging traditionally underserved students in scientific inquiry, though measures of achievement varied depending on the assessment method used (Ketelhut et al., 2010). Research on Quest Atlantis has shown that game players’ behavior changes from the game also transferred to their behaviors outside the game world (Barab et al, 2007).

In online discussion environments modeled as communities of inquiry in higher education and professional communities, researchers found it difficult to observe the more sophisticated phases of inquiry analogous to analysis and theory-building (Garrison et al., 2005). Such scientific inquiry practices are evident, however, in free-choice digital gaming environments such as, World of Warcraft (WoW), which engages millions around the world. Researchers found that it is not unusual for players to gather data in spreadsheets, create models of the data in the form of simple mathematical equations, and then argue about whose model was “better” in terms of prediction and explanatory scope (Steinkeuhler & Duncan, 2008). This research may suggest that either gaming may foster deeper levels of scientific inquiry and/or it may be easier to observe and measure inquiry in gaming environments.
3. Models of Scientific Inquiry

Many models of scientific inquiry are rooted in a theory of argumentation from Toulmin (1958) that centers on claims, data or evidence, and warrants, which are principles we hold to be true. Kuhn (2005) describes argumentation as the coordination of theory and evidence, which scientists carry out with conscious control and explicit and consistent criteria. In developing understanding, learners must also test ideas against evidence and continuously revise theories (Harlen, 2005). This process is inherently collaborative as players must compare and contrast their claims and evidence with those of others to advance the knowledge building process (Scardamalia & Bereiter, 1996).

4. Design and Implementation of Martian Boneyards

EdGE created the land of Arcadia in the MMO, Blue Mars, to embed games of scientific mystery (see Figure 1). Arcadia is a city² in Blue Mars that contains an abandoned Science Center and extensive grounds around the Center. Martian Boneyards is a prototype game with a mystery storyline that takes place in Arcadia.

Figure 1 goes about here.

4.1 The Storyline of the Martian Boneyards

Throughout the game, the storyline evolved incorporating the players’ ideas and evidence. The designers initiated the storyline: A team of explorers found the Science Center abandoned and containing semi-functional research tools from previous settlers. They also found a gruesome discovery–there were bones scattered all over the
surrounding outside area. That is when the explorers knew they had to call in a group of players to help them figure out what had happened in *Arcadia*. From this basic premise, the designers allowed the storyline to grow in the directions pursued by the players.

### 4.2 Inquiry Tools

The tools designed to support players’ inquiry naturally fell into four categories of scientific inquiry: exploration, data-gathering, analysis, and theory-building.

When players first came to *Arcadia*, they would take some time to explore the area. They might click on artifacts and look for clues about what is happening in the Science Center. They would see scientific posters on the walls and skull models in the display cases.

Within their first session into the game, most players became engaged in data gathering—actively hunting bones with their PDA (see figure 2). This required players to move around the Boneyards, zooming in to scan for very small pieces of bones sometimes obscured by bushes or sand. Designers strived to make discovery of artifacts simple enough to not be frustrating or lose players but also challenging enough to keep players engaged.

Figure 2 goes about here.

Players shared and analyzed their data by clicking on analysis workstations in the Science Center (see figure 3). The workstations allowed players to see two views of most artifacts, take and record two linear measurements (i.e. length and width) for each view and record notes and tags for other players to see. An artifact would only show up on a
player’s table if they, themselves, had found it. This was meant to inspire players to work together to make sure everyone found all the artifacts. Collaboration was also built into a verification process where an artifact had to be found by a certain number of people (originally 20) before it would be verified and available for use as evidence for a proposed claim on the theory-building board. This constraint was meant to foster collaboration, slow-the pace of early players to allow for the community to grow, and also mirror a peer-review process in science where data are not taken seriously until replicated in some manner by colleagues. Similarly, the instructions for using the measurement tool were deliberately left wide open to see if the community would identify the need for standardization and attempt to come to agreement on procedures.

Figure 3 goes about here.

A theory-building board was designed for players to post claims about the artifacts they were finding, what may be in the Boneyards, and how it got there (see figure 4). It allowed players to add, edit, and publish their own claims and comment on the claims of others. When posting a claim, a player was required to include evidence, in an attempt to foster evidence-based inquiry in the game. To do this, the player posted an image of the artifact(s) they were citing, and could then describe the related tags or measurements, as part of their overall claim. Other players could comment on the claim, adding their own evidence to counter or confirm the original posting. There was also a peer-review area where, after an author felt the claim was ready to publish, the community could rate the claim.
4.3 Advancement Structure

Gamers are typically motivated by some sort of advancement structure in a game (Rollings & Morris, 2004; Yee, 2009). EdGE used an informal reward system that relied on the designers’ characters to give out awards to recognize players’ activities. The nimbleness of this approach allowed designers to respond to the players’ stated desires using observations to identify what seemed to motivate productive game play. EdGE commissioned a clothes designer in *Blue Mars* to make t-shirts, cargo pants and vest (with lots of pockets for collecting bones), and a water bottle—the type of reward the players’ requested and all items one might need on an archeology dig.

5. Research Study

The research of the design and implementation of *Martian Boneyards* uses mixed methods to examine:

1. What types of design elements and implementation strategies did EdGE designers use to support sustained scientific inquiry in *Martian Boneyards*?
2. Who came to play *Martian Boneyards*? Who became involved in sustained scientific inquiry in the game?
3. What is the nature and quality of players’ scientific inquiry in *Martian Boneyards*?

5.1 Methods

This research study uses methods borrowing from ethnography to study the context within which players act and from phenomenology to study how players’
experience the gaming environment. This multiple viewpoint addresses the interdependency of environment design, players’ activity, and players’ progress by representing the game as a distributed system of players, knowledge, and scientific tools and resources in *Arcadia*.

Netnography, also referred to as virtual ethnography, blends analysis of digital records from Internet-supported interactions with ethnographic methods used when the researcher is immersed in the community of study (Kozinets, 2002; Hine, 2000). Netnographic methods used in this research include analysis of electronic records of avatar motions and actions as well as surveys, interviews, and participant observations to provide a well-rounded picture of the behaviors and culture of the participants in the environment.

### 5.2 Sample

The overall player population for this study is all entrants (N=613) to *Arcadia*, which includes anyone who took the registration survey between June 1, 2010 and Sept 30, 2010. Entrants must have stated that they are 18 years or older and consent to be “keytracked” (anonymously monitored for clicks on *Arcadia* tools) to be included in the sample.

The player community was recruited primarily from the current beta-test community of *Blue Mars* so it waas primarily composed of experienced virtual world users who were very dedicated to exploring the *Blue Mars* environment and were seeking activity in the new virtual world.

The demographics of the sample are reported in Tables 1 and 2 in terms of 3 nested sub-samples. *All Players* (N=228) are those who interacted at least once with an
inquiry tool in *Arcadia*. **Core Players** (N=66) are those who used the inquiry tools > 20 times and **Top Tier Players** (N=18) are those who used the inquiry tools > 100 times.

The typical *Martian Boneyards* entrant and player is a 36-year old white male who is not involved in a science career but may read or watch TV about science. He spends a lot of time in virtual worlds. The only notable difference in demographics from looking through the sub-samples is that women make up a larger part of the sample when filtered down to the top tier. The player base also becomes more predominantly white as it narrows (see Table 1).

5.3 Measures

The design and implementation study examined the relationship between game design, players’ activity, and players’ progress in *Martian Boneyards*.

5.3.1 Game Design

This research focuses on the decisions and facilitation strategies used by the designers intended to support and sustain scientific inquiry. These decisions include choosing the scientific resources, crafting collaboration tools, conceiving a storyline, and placing all of this in an environment for the game play.

5.3.2 Players’ Inquiry Activity

Players’ scientific inquiry activity in *Arcadia* is measured two ways. The frequency is measured of the interactions with each of the inquiry tools. The number of scans each player conducts with the PDA is a measure of their data gathering activity. The extent of players’ analysis activity is recorded as the number of their interactions (tagging/measuring/comparing) with the analysis workstations, and their extent of theory-
building is number of interactions with the theory-building board. This numerical measure of theory-building is actually an underestimate because it only counts the theory-building activity that took place within Arcadia, excluding the substantive activity on the web-based discussion board that players chose to use outside the gaming environment.

The second measure of extent of inquiry is players’ overall duration of time spent in Arcadia. The avatar activity logs recorded each avatar’s entry and exit into each room or outside area in Arcadia. To avoid including the idle players in the research data, the records that were greater than one hour in duration were removed if there was no other activity recorded by that avatar subsequent to the one hour.

5.3.3 Quality of Scientific Inquiry

A team of three scientists who specialized in paleo-anthropology and biology reviewed a set of user-generated materials that represented their knowledge building in the game using a rubric and process modified from previous research to review the quality of materials online science courses (Rowe & Asbell-Clarke, 2008). The Martian Boneyards materials included postings from the theory-building board (with supporting evidence), postings from the Blue Mars web forum that players used for supplemental discussion, and one excerpt from one in-game chat (over 200 text entries in total).

The materials reviewed included all the entries relevant to scientific content and excluded posts that were purely social, or purely about the storyline. One inquiry topic, which was one of the most interesting, (the size of the Baobab tree) was only addressed in chat and therefore the chat log from this particular discussion was included in the materials.

The panel of scientists used a rubric to rate:
a) The extent of the scientific inquiry
b) The sophistication of the scientific inquiry
c) The accuracy of core ideas in comparative anatomy
d) The depth of core ideas in comparative anatomy

They rated the quality of the entire set of materials along each dimension on a 5-point scale (poor, fair, good, very good, excellent) discussion related to a project conducted in an introductory undergraduate class for non-science majors.

The scientists agreed that comparative anatomy was the core area that was covered in most of the materials. The materials touched on other topics such as evolution, genetics, and botany but the sustained inquiry was in the area of comparative anatomy.

They were in agreement that their ratings representative the overall quality of the materials and noted that there were very few exceptions to the accuracy of content ratings or quality of the science information resources used in the game. The variability in the ratings across the set of materials was primarily in the extent of inquiry in the natural ebb and flow of the discussion.

5.4 Data Sources

The netnographic techniques used by the researchers examine a variety of digital sources to get a broad and deep look at the context that is mediating the game play, the environment and the community, and the player’s experience within that context. The data sources include surveys, digital records, observations, and interviews.

5.4.1 Surveys

Players were required to take an initial survey at the entrance to Martian Boneyards. When providing consent for access, entrants were asked their sex, race, age,
science involvement, and virtual-world experience. All survey responses were tagged with an anonymous ID allowing them to be linked together with the other digital data collected.

5.4.2 Avatar Log

Each time a player clicked on any tool in the Science Center, the interaction was recorded with a time stamp and the anonymous player ID. Each tool was associated with one phase of the inquiry cycle: data gathering, analysis, or theory-building.

5.4.3 Participant observations

The team of three designers also played characters in the game. As a result they were able to serve as participant observers. The participant observers recorded their own actions during each 2-hour event (at least twice per week), as well as recorded the tone and events of the players’ activities during events and any non-event times they monitored during each week. Observations focused on storyline, social dynamics, and how designers supported scientific inquiry. Two additional members of the research team, who were not designers or regular participants in the game, co-observed and reviewed one summary event report for validation.

5.4.4 Design Documents

Designers (EdGE and VSE) recorded their decisions throughout the process and made these available to the research team. The design documents focus on the scaffolding of evidence-based research in the design of the inquiry tools, embedding scientific inquiry within an evolving storyline, and making the environment as attractive and immersive as possible with limited funds.
5.4.5 Interviews

Researchers used a semi-structured interview protocol to conduct avatar-to-avatar interviews, situating the interview in the studied environment (Turkle, 2005). The interviewer was an assistant researcher and project coordinator for EdGE. Players were asked what attracted them to Arcadia, what they felt contributed to the value of the game for enjoyment and science learning, and how their experience in Martian Boneyards changed the way they think about science.

6. Results

The results from the Martian Boneyards implementation study are reported in terms of the game design, the player activity, and player progress.

6.1 Game Design for Martian Boneyards

The game design elements for Martian Boneyards focused on scaffolding and sustaining collaborative inquiry among the player community. The team constantly sought a balance in how to deliver a compelling and evolving storyline, how much instructional design to include in the game, and how to help the players organize their evolving ideas while also providing a community-centered entertaining game.

The game started on Jun 1, 2010 and ended October 7, 2010 and was implemented in 5 phases. Each phase corresponds to the opening of a new area of the Boneyards. Phase one lasted about 7 weeks, while the other four phases were each closer to 3 weeks. The staged roll-out of Arcadia allowed designers to schedule the opening of each new phase at the right time for the community. Designer characters were able to roll the timing of the phase releases into the storyline as well, using them as rewards for
the good work the community is doing—players were told that the center could open the new section because they got more funds for from a mysterious donor who wants to spur them on to get more answers to the mystery.

The process of changing the actual *Blue Mars* environment was costly and also had a long “build” process between the designers and when it was tested and integrated into the MMO. This made “on-the-fly” changes to the environment very limited. The designers created screens in the Science Center that enabled streaming of a flash interactive from the web so that they could create a simple web page to display in *Arcadia* permanently, even when the designers’ characters were not there. This tool, referred to by designers as a flashboard, became essential for posting updates. For example, designers posted temporary fixes for any known problems with tools that could be patched until it was resolved when the next build was released. Throughout the game, the flashboard were also used to communicate basic instructions on how to use tools, storyline components (e.g. essential questions for players to focus on), and posters of scientists using a comparative approach to studying bones.

Chat was the primary communication tool in the game. A chat window was present in the lower part of a players’ screen and displayed all the public chat that was occurring by all players in *Arcadia*. This meant that the conversations of players in the Boneyards were intermingled with conversations of players in the analysis room or elsewhere, and sometimes they were even in different languages, so the dialogue at times became hard to read. Players who wanted to chat one-to-one could move to a private chat with their identified “friends”, but observers and others could not see this chat.

6.1.1 Storyline
Though players and designers agreed that the storyline was possibly the most important element for the success of the game, ironically it was one of the least formed elements of the design. The initial storyline presented to the players in the poster in the *Arcadia* entrance hall was very vague and basically only stated that Laurel, Fischer, and Tieaun (the designers’ characters) had discovered the science center, along with the surrounding Boneyards, and needed the community’s help to figure out what had happened before they had arrived. While the artifacts were placed with a very intricate backstory in mind, the designers in the game did nothing to direct the play toward that specific interpretation of the evidence.

It was clear very early to the designers that intrigue and mystery were paramount to several of the players. While everyone enjoyed the quest of hunting, those who came back night after night wanted more information about the backstory. Top players, mostly female, asked many questions about the (fictional) lives of the explorer characters and they were quick to pick up on any (often spurious) piece of information that characters dropped that could be taken as a clue.

For example, on June 30 in the Boneyards, one player, Laurel, remarked—with no particular goal in mind—that the bones seemed hard to see in the river when the sand partially buried them. That simple comment provoked the players to form into a search party to comb the river. They spread out in a line and went row by row over the grounds to search for bones (see figure 5).

---

Figure 5 goes about here
6.1.1.1 Character Development
The designers felt that the relationship that players developed with each other and with the characters designers was a major reason players came back night after night. Some of the characters, such as the three explorers, were known from the outset of the game. Other characters, also played by designers, were lurking in the Science Center occasionally and emerged secretly to have private conversations with players so that designers could watch what happened to the “secrets” that they passed on. A third set of characters was very important to the storyline, those were not represented by avatars, but only by bones and skeletons in the Boneyards. A key non-avatar character in the game was the third skeleton, JJ Cleat.

6.1.1.2 JJ’s Journal
While players were translating the posters, the designers recognized that they needed to acknowledge this vast amount of effort. The designers accommodated this need by providing some hints to a closure to the story through the introduction of a journal from one of the missing scientists, JJ Cleat. The journal could be opened by a clue in the posters.

JJ’s journal consisted of no more than a series of postings on the web-based discussion forum that players used for communication, but was absolutely essential for providing closure to the game experience for players. The designers didn’t write the journal until the players had gotten far enough along to determine a rough storyline of their own so that the journal could be consistent with the players’ evidence and reasoning. The creative writing style of the journal gave just enough information to answer players’ questions about happened while also leaving room for their imaginations.
6.1.1.3 Award Ceremony

Late in the game, the designers realized that since they wanted to bring the game to conclusion, the storyline needed to enforce a deadline. Midway through the game, Laurel had fortuitously introduced the concept of an interplanetary council spontaneously and the designers built on this so that it rippled through the final month and a half of the game. It gave designers and players a focus to achieve storyline closure. It also gave designers a way to prompt the players towards a synthesizing phase of inquiry. The promise of a community-based award was all that was offered to players, though when the ceremony took place, individual awards were also given. The awards night gathered the largest simultaneous crowd ever in Arcadia (about 40 players) and provoked much excitement in the lead-up and for the evening itself (see figure 6).

Figure 6 goes about here

6.1.2 Facilitation

The storyline was an ongoing discussion among the designers as they played their characters and responded to the ideas and activities of the players. The role of designer character enabled the designers to have unique opportunities for facilitation and participant observations. It was a perfectly natural part of the storyline for the designers’ characters to be asking “Why do you think that? What is your evidence?” when a player offered an idea—both scaffolding the players’ articulation of their evidence-based arguments, and also prompting impromptu “think-alouds” enabling researchers to see the players’ explanations of their scientific thinking in text as they respond in the chat box.
6.1.3 Collaboration Tools

Collaboration was a main driver for the design of the scientific inquiry tools. Once players mastered the use of the PDA, they trained new players on the correct and most efficient way to use it. In the beginning, it was the designers who gave the players tours of Arcadia and the tools, but that task was quickly taken over by other players.

The theory-building board required that all claims be posted with supporting evidence from the workstations. This in-game tool was not functional very early in the game and was never the medium of choice for players who wanted to share text along with images and URLs to Internet resources (this was still awkward to do in Blue Mars). Players used a web-forum (hosted for the Blue Mars community) to share information about the game. Some players posted their entire inventory of artifacts, along with descriptions of where they were found, to help other players.

Participant observers noted that players quickly adopted the “rules” of the inquiry tools, and the scientific language and behaviors promoted by the game were enculturated within the community. Players were noted as saying phrases such as “we can’t use it as evidence until it has been verified”, “before we can post a claim we have to find evidence to support it”. Players were using this language as one might explain the rules of a game to a new player, but later speaking about the need for evidence to substantiate theory-building became commonplace in the community. Even as players moved to using web-based theory boards, they still used language suggesting that they were trying to form evidence-based arguments.
6.2 Players’ Activity

The players’ activities are described in terms of the extent and nature of their interactions with the collaboration tools and scientific resources.

6.2.1 Extent of Scientific Inquiry: Quantitative Results

The extent of the scientific inquiry is measured by the frequency of interactions with inquiry tools and duration of play in the game.

6.2.1.1 Frequency of Interactions with Inquiry Tools

Table 3 shows the frequency of inquiry tools used by core players in the game, overall and disaggregated by participant characteristics. Overall, approximately 74% of players’ interactions were in the data-gathering phase of inquiry, which is consistent with players’ own statements during think-alouds that they like quests and hunting in virtual worlds. Interestingly though, 15% of the interactions were in analysis and another 11% were using the theory-building tools. These phases are similar to the synthesis and resolutions phases of the Community of Inquiry model used to study inquiry in online educational environments (Garrison et al. 2003). Those researchers, however, found inquiry lacking in these later phases in online communities of inquiry (Garrison et al., 2005). This finding suggests that MMO games such as Martian Boneyards might be better at scaffolding inquiry in these phases and/or that netnographic methods are more able to capture this type of inquiry.

Figure 7 shows a histogram of the frequency of inquiry tools used by core players in the game, for everyone and disaggregated by males and females. The average participation with the inquiry tools is higher for females than males in the core group.
(F(1, 60)=5.209 , p<0.05). Females also have higher participation in analysis (F(1, 60)=6.156 , p<0.05) and theory building (F(1, 60)=4.674 , p<0.05) activities. The only inquiry phase where activity is not significantly different between males and females is in data gathering.

Figure 7 goes about here

6.2.1.2 Duration of Play

Table 4 shows the duration of play, in hours, of players, overall and disaggregated by player characteristics. The 66 core players spent an average of nearly 28 hours in Martian Boneyards. A large extent of players spent less than 20 hours in Arcadia, but some players spent upwards of 200 hours in the game. Because of the skew, particularly among the top tier players, the apparent large difference in means between males in females is not significant (See Figure 8 for histogram).
6.2.2 Nature of the Community’s Scientific Inquiry

Observers noted the inquiry among players was highly collaborative, and overall, the activity within the game was highly on task. The inquiry is described here as four phases: exploration, data gathering, analysis, and theory-building, though it is important to remember that these are not so much separate phases—and are certainly not linear—when they take place in the game.

6.2.2.1 Data Gathering

Players were eager to have everyone in the community locating artifacts and contributing their findings. The players clearly felt a sense of community and enjoyed helping each other. They also depended on others to scan and “verify” artifacts before those artifacts were considered valid evidence. Leaders emerged in the group early whom the designers were able to rely upon to help facilitate the game. At no time was any player observed trying to hide their inventory or not readily sharing their data.

6.2.2.2 Analysis

Individual players tagged the bones, and there was not as much discussion as anticipated about the tags themselves. There was, however, substantive collaboration around the measurement of bones, as that became important in the players’ identification of the bones. As players continued to discuss how they measured bones, it became clear that there was no standard amongst the players. The designers suggested that people write about how they measured each bone in the comments section of their workstation, as this was designed to be shared information. Players chatted about standardizing
measurement but did not use the tools to create a formalized system. For example, observers noted:

Francie brought up how to tell male from female again re: using pelvis as well as ratios of bones….Does that make it a male shoe or female shoe, consensus was that it is too big for female, therefore Scully is male. They are linking shoe and bones by shoe to Scully, that is good but they need to do some more comparison/measurements to link the bones to each other via a theory…. Francie thought that you multiplied bone length by some number (2.6?) to get height but she wasn’t sure.

At this point, other players had started identifying different species that the bones might be from. They brought in resources from the Internet, their own background knowledge, and sometimes real-life research trips were suggested. A player, Kalw, posted on the discussion board:

Hi! We have come a long way baby! Ok now I do not know if you guys have found this site or not it is http://www.whyevolution.com/chimps.html and I forgot all about Pittsburgh's Carnegie museum of Natural History!!! May take a road trip also )) You would not believe what they have in that museum!! Oh Notail’s noticed on the sign in the work station area that it says btw jj - checkout arcadiashare.terc.edu did it always say that about jj ? love and hugs kalw

To which another player, Jespau, replied:

Hi kalw -

Nice work on chimps. I spent ages today researching them after your idea and
you are right. They are our closest primate relative. Only 1 chromosome different. That’s what’s on the painting i think. - Jespau

Soon the measurements became very detailed and players began to compare their measurements to outside information. Jespau posted the following:

[Note: The headings B-11-XXX refer to the artifact IDs used by the game tools.]

**FEMUR: Thigh bone**

The average adult male femur is 48 centimeters (18.9 in) in length and 2.34 cm (0.92 in) in diameter. This is the longest bone in the human body and a quarter of the body length. (Brothwell 1981: 35). 5 Femur bones have been found. A persons height can be calculated from these measurements.

These two bones match in size and could be from the same living thing. This is short for a Mature HUMAN but could still be a younger person.

- **B-11-HQ1**: 35.82 cms x 8.55 cms (width) 2.73 (mid width) (14.10 inches x 1.07 inches)
- **B-11-UE1**: 35.82 cms x 8.73 cms (width) 2.91 (mid width) (14.10 inches x 1.14 inches)

Colour: Beige and the other a bit more sandy but close match.

Height: 143.28 cms or 4 feet 7 inches.

Only one found. To short for human. ANIMAL?

- **B-11-MK1**: 17.22 cms x 3.04 cms (mid width) (6.77 inches x 1.19 inches)

Colour: Light Tan

Height: 68.88 cms or 2 feet 3 and three quarter inches.
These two bones match closely in size, are close to HUMAN length and
could be from the same HUMAN of short stature.

B-12-DW1: 41.36 cms x 2.64 (mid width) (16.28 inches x 1.03 inches)
B-12-QG1: 41.64 cms x 2.27 (mid width) (16.39 inches x 0.89 inches)

Height: (averaged) 166 cms or 5 feet 5 inches.

Colour: DW is ancient looking and white with tan blotches. QG1 is more
grey with no tan blotches. So not a great match for colour but both
found in Cave 12. These last two bones are not verified yet so no
picture evidence available. See your workstation for examples.

6.2.2.3 Theory Building
The term “theory-building” was controversial throughout the design of Martian
Boneyards. Originally the tool designed for Arcadia was called the theory board. An
advisor to the project expressed concern that the name could promote the misconception
that scientific theories are guesses (i.e. “it’s just a theory”). Understanding this, but
wanted to stick to a model of the coordination of evidence and theory, the team settled on
the name theory-building board with the posts being called claims.

The players never converged on one substantiated set of claims that could emerge
into one theory. Rather, they had various related and overlapping claims, some having
much more detail (depending mostly on time and imagination expended by the player)
than others.

6.3 Player Progress
The quality of the scientific knowledge building was measured through a review
of player-generated artifacts by a panel of scientists. The artifacts reviewed include all
claims posted on the theory-building board (with supporting evidence) along with postings on the *Blue Mars* web forum that players used for supplemental discussion and one excerpt from one in-game chat on a topic not well represented in archived postings (over 200 text entries in total). The team of three scientists concluded that substantive scientific inquiry took place in the game.

The player community engaged in scientific inquiry—questions, making claims, substantiating claims with evidence—to an extent that would be considered very good in an undergraduate introductory science course. The content generated in comparative anatomy was rated very good on accuracy and good on depth. Reviewers noted that the game motivated a level of inquiry among some players that was similar to top students in a class who took a lesson much farther than required out of self interest. One reviewer commented “Those top players reminded me of those students you get once in a while that just have a burning desire to learn”.

Reviewers also noted that most, if not all, Internet resources used by players were from reasonable scientific websites, including Wikipedia and accredited sites from universities and national labs). There were no personal blogs or websites from non-scientific interest groups found. Players had a grasp of how to use the material found, the only criticism in the reviewers’ minds was that many of related topics were dealt with superficially (e.g. players did not dig deeply into evolution or functional morphology, though there was conversation leading in those directions). Reviewers agreed that nearly all of the content in comparative anatomy was accurate and players’ arguments in these areas were scientifically valid.

7. Discussion
Martian Boneyards was designed to create a community of inquiry that comes together to solve a mystery. EdGE strove to create authentic open-ended inquiry for the players by having players’ contributions to the scientific inquiry judged by their validity according to the evidence available, not against “right” or “wrong” answers known by some external authority.

A question that arose repeatedly in the study of Martian Boneyards was the amount of instructional scaffolding needed for successful gameplay and for high quality scientific inquiry. EdGE designers hypothesized that players would need very little direct instruction (teaching of science content) and thus provided none, and indeed, researchers found that players were able to produce high quality content and methods on their own. EdGE also developed minimal instructional scaffolds to help game play (teaching of tools and storyline), primarily using designers who were playing characters in the game to give storyline updates to players and let the information diffuse “naturally” through the game networks. There is some evidence that this method was successful in that key players gained status in the game for being able to fill in others on the storyline and game mechanics, as well as the scientific knowledge being developed. On the other hand, many more players may have stayed and engaged with the game had there been more instructions on how to get started.

EdGE comes away from this research even more convinced that minimal instructional design is appropriate for a game, in terms of text or even video-type tutorials; however, the user interfaces for the tools need to be highly intuitive to not require instruction. Today’s gamers have many accepted norms of where buttons should
appear and how interfaces are expected to work. These should guide the user interface as much as possible so that tools behave the way a player expects.

The Martian Boneyards players frequently brought in resources from the Internet to help the community solve the game. In fact, they were frustrated by the limitations of being unable to display external websites within the Science Center. Integration of these media would have allowed this audience to use the Internet as their “library” in the Science Center. The links that players posted on the discussion board were from scientific cites and rated as suitable by the expert reviewers. This leads EdGE to believe that not much design effort needs to go into finding and filtering appropriate content for players’ investigations (as long as the designers know quality material is readily available on the subjects at hand). Instead, time should be spent on the user interface of the tools, the attractiveness of the environment, and ensuring that the storyline yields many opportunities for high quality scientific inquiry.

EdGE found that designers’ participation as characters within the game and their ability to provide a dynamic and responsive storyline helped support this participative culture and promote independent high-quality scientific knowledge building in Martian Boneyards. The community was also central to shaping the experience and the environment. These finding have strong implications for issues many issues that will be important for learning environments of the future.

Martian Boneyards is an example of an emergent participatory culture where the public is taking control of their own learning (Falk & Dierking, 2010; Jenkins, 2006). This raises questions about the role of designers in fostering players’ scientific inquiry, how to design gaming environments that are sustainable both from a gameplay and a
financial standpoint, and how to leverage the emergent identities and the decentralization of knowledge building in a participatory gaming community. The participatory model invokes decision-making and knowledge building using crowdsourcing and wisdom of crowds. These models are interesting because they decentralize authority and provide no clear source of authority of knowledge. This runs very counter to current educational hierarchies in academia.

We are approaching a time when the high school and university students, and then the first wave of entrance to the professional workforce are learners who grew up in a digital world—they are the digital natives. These are likely very different learners than their parents and teachers, and they will likely be very different workers than the managers who hire them. This new generation thinks in social networks, and their lives are distributed across media and across social and learning contexts in ways educators never previously could have imagined.

The study of Martian Boneyards is a first step in understanding how social digital gaming environments can be used to transform how we think about learning environments. A highly immersive and aesthetically MMO environment with a carefully-crafted storyline and set of collaboration tools was able to leverage gamers’ desires to solve problems and work collaboratively and engaged them in sustained collaborative scientific inquiry.

The research on the design and implementation of Martian Boneyards adds to the growing evidence that gaming environments might be harnessed, along with gamers’ vast amounts of time, energy, and “blissful productivity”, to foster productive scientific
inquiry in innovative learning environments of the future. EdGE will continue on the quest to make this happen.
References


Galas, C., & Ketelhut, D. J. (2006). River City, the MUVE. *Learning and Leading with Technology, 33*(7), 31-32.


Acknowledgements

The research on *Martian Boneyards* that helped frame this work and a portion of the writing of this paper was supported by a grant of the National Science Foundation (NSF-0917520) to the first two authors. In addition, the authors are grateful to TERC for additional funding to support new directions in the research. We wish to also thank Herve Gomez and Reed Knight our fabulous game designers and the EdGE advisory board, in particular Diane Ketelhut, Brian Nelson, and Chris Dede, for their wisdom to help guide our study.
Footnotes

1 TERC, formerly known as the Technical Education Research Center, is a not-for-profit research and development organization focusing on technology-based math and science education.

2. Each region rented and developed on Blue Mars is called a city. Arcadia is a city and has its own entry point on the Blue Mars launch page. All Martian Boneyards activity happened within the city, Arcadia.
Table 1

*Sex, Race, and Age of Martian Boneyards Players*

<table>
<thead>
<tr>
<th></th>
<th>All entrants (N=613)</th>
<th>All players (N=228)</th>
<th>Core players (N=66)</th>
<th>Top players (N=18)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>29%</td>
<td>29%</td>
<td>32%</td>
<td>50%</td>
</tr>
<tr>
<td>Male</td>
<td>66%</td>
<td>66%</td>
<td>60%</td>
<td>50%</td>
</tr>
<tr>
<td>No Answer</td>
<td>5%</td>
<td>6%</td>
<td>8%</td>
<td>0%</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>74%</td>
<td>78%</td>
<td>76%</td>
<td>83%</td>
</tr>
<tr>
<td>Non-white</td>
<td>26%</td>
<td>22%</td>
<td>24%</td>
<td>17%</td>
</tr>
<tr>
<td><strong>Age (yrs)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-27</td>
<td>31%</td>
<td>32%</td>
<td>35%</td>
<td>39%</td>
</tr>
<tr>
<td>28-37</td>
<td>27%</td>
<td>27%</td>
<td>27%</td>
<td>33%</td>
</tr>
<tr>
<td>38-47</td>
<td>23%</td>
<td>20%</td>
<td>15%</td>
<td>11%</td>
</tr>
<tr>
<td>48-57</td>
<td>13%</td>
<td>13%</td>
<td>12%</td>
<td>11%</td>
</tr>
<tr>
<td>58-67</td>
<td>5%</td>
<td>7%</td>
<td>9%</td>
<td>6%</td>
</tr>
<tr>
<td>68+</td>
<td>1%</td>
<td>1%</td>
<td>2%</td>
<td>0%</td>
</tr>
<tr>
<td><strong>Mean age</strong></td>
<td>36.00</td>
<td>36.16</td>
<td>35.71</td>
<td>33.00</td>
</tr>
</tbody>
</table>

*Note. Core players had >20 tool interactions, top players had >100 tool interactions*
Table 2

*Science and Virtual World Interest of Martian Boneyards Players*

<table>
<thead>
<tr>
<th></th>
<th>All entrants (N=613)</th>
<th>All players (N=228)</th>
<th>Core players (N=66)</th>
<th>Top players (N=18)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Science in daily life</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not involved</td>
<td>26%</td>
<td>22%</td>
<td>33%</td>
<td>28%</td>
</tr>
<tr>
<td>Interested</td>
<td>48%</td>
<td>51%</td>
<td>44%</td>
<td>44%</td>
</tr>
<tr>
<td>Involved</td>
<td>25%</td>
<td>27%</td>
<td>23%</td>
<td>28%</td>
</tr>
<tr>
<td><strong>Time spent in virtual world</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nearly all</td>
<td>12%</td>
<td>9%</td>
<td>9%</td>
<td>0%</td>
</tr>
<tr>
<td>A lot</td>
<td>55%</td>
<td>58%</td>
<td>54%</td>
<td>50%</td>
</tr>
<tr>
<td>A little</td>
<td>20%</td>
<td>21%</td>
<td>26%</td>
<td>39%</td>
</tr>
<tr>
<td>First time</td>
<td>14%</td>
<td>12%</td>
<td>11%</td>
<td>11%</td>
</tr>
</tbody>
</table>

*Note.* Core players had >20 tool interactions, top players had >100 tool interactions.
Table 3

*Distribution of Inquiry Phases by Demographics*

<table>
<thead>
<tr>
<th>Player Variable</th>
<th>Total tools</th>
<th>Data gathering</th>
<th>Analysis</th>
<th>Theory building</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>118.27</td>
<td>87.00</td>
<td>17.95</td>
<td>13.32</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male (n=41)</td>
<td>91.78*</td>
<td>71.41</td>
<td>12.49*</td>
<td>7.88*</td>
</tr>
<tr>
<td>Female (n=21)</td>
<td>170.00*</td>
<td>117.43</td>
<td>28.62*</td>
<td>23.95*</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-white (n=13)</td>
<td>79.31</td>
<td>68.69</td>
<td>6.69</td>
<td>3.92</td>
</tr>
<tr>
<td>White (n=49)</td>
<td>128.61</td>
<td>91.86</td>
<td>20.94</td>
<td>15.82</td>
</tr>
<tr>
<td><strong>Science in Daily Life</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-career (n=48)</td>
<td>123.42</td>
<td>94.60</td>
<td>15.94</td>
<td>12.88</td>
</tr>
<tr>
<td>Career (n=14)</td>
<td>100.64</td>
<td>60.93</td>
<td>24.86</td>
<td>14.86</td>
</tr>
<tr>
<td><strong>Time spent in Virtual Worlds</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high (n=42)</td>
<td>98.38</td>
<td>71.83</td>
<td>15.02</td>
<td>11.52</td>
</tr>
<tr>
<td>low (n=20)</td>
<td>160.05</td>
<td>118.85</td>
<td>24.10</td>
<td>17.10</td>
</tr>
</tbody>
</table>

*Note.* N=62 because the 4 players who chose not to identify sex are not included in analyses.
*p < .05*
Table 4.

*Time Spent in Game by Core Player Types (N=66)*

<table>
<thead>
<tr>
<th>Player Type</th>
<th>Mean Duration (hrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>27.68</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
</tr>
<tr>
<td>Male (n=40)</td>
<td>20.80</td>
</tr>
<tr>
<td>Female (n=21)</td>
<td>40.78</td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
</tr>
<tr>
<td>Nonwhite (n=13)</td>
<td>13.70</td>
</tr>
<tr>
<td>White (n=49)</td>
<td>30.15</td>
</tr>
<tr>
<td><strong>Science in Daily Life</strong></td>
<td></td>
</tr>
<tr>
<td>Non-career (n=51)</td>
<td>26.19</td>
</tr>
<tr>
<td>Career (n=15)</td>
<td>26.00</td>
</tr>
<tr>
<td><strong>Time spent in Virtual Worlds</strong></td>
<td></td>
</tr>
<tr>
<td>high (n=42)</td>
<td>23.28</td>
</tr>
<tr>
<td>low (n=24)</td>
<td>31.20</td>
</tr>
</tbody>
</table>
Figure Captions

Figure 1: The Land of Arcadia in the MMO Blue Mars

Figure 2: A player uses the PDA to collect data in Martian Boneyards

Figure 3: Players using the workstations that allowed measurement and comparison of artifacts

Figure 4: Theory board where players post evidence-based claims and reasoning

Figure 5: Players Organize in a Search Party to Hunt for Bones

Figure 6: Players were fully decked out in eveningwear for the awards ceremony

Figure 7: Histogram of Frequency of Inquiry Tools Used by Core Players, Total and By Sex

Figure 8: Histogram of Duration of Play, Overall and by Sex
A player finds a bone under a bush and uses the PDA to scan it.

The PDA shows the scanned bone and allows the player to enter information about it.

The browse mode allows the player to see all the bones already scanned.
Players use the workstations in the science center to share and analyze data.

The tools enable players to measure the bones for each view (left) and add that data, along with tags for name, species, and other features. An ongoing ranking system allows players to see the tags used by others in order of their frequency.
Martian Boneyards Figure 7

[Diagram showing bar charts comparing total tools, data gathering, analysis, and theory building for overall, male, and female categories.]