2006-2132: ENHANCING ENGINEERING OUTREACH WITH INTERACTIVE GAME ASSESSMENT

Leilah Lyons, University of Michigan
Leilah Lyons is a doctoral student in the Dept. of Electrical Engineering and Computer Science at the University of Michigan. Her interests include use of technology in informal and pre-college education.

Zbigniew Pasek, University of Windsor
Dr. Pasek is an Associate Professor at the Dept. of Industrial and Manufacturing Systems Engineering at the University of Windsor, Canada. He was previously with University of Michigan. His research interests include manufacturing automation and system design, informal engineering education, and decision-making processes in organizations.

Enhancing Engineering Outreach with Interactive Game Assessment

Abstract

The need to educate general public about technology grows with broadening gap between technology use and its understanding in a consumer society. One of the effective venues for such education is that of hands-on museums, which engage a wide spectrum of visitors. This paper reports on the use of a data collection mechanism embedded in an interactive museum exhibit that highlights principles of modern consumer product engineering. The exhibit is comprised of a set of computer games and complementary physical displays. The games have a built-in data collection system tracking users' actions while playing the game. Collected data allows for demographic analysis of visitor population, user performance assessment, and provides game-play perspective useful for effective game design. Presented results are based on a year-long study involving about 17,000 museum visitors.

1.0 Outreach in the Form of a Museum Exhibit: Overview of the Project

Government funding supports research work on the cutting edge of manufacturing technologies, but the general population's understanding of manufacturing processes, equipment, and careers lags far behind that edge. To bridge the gap, the NSF Engineering Research Center for reconfigurable Manufacturing Systems (ERC/RMS) at the University of Michigan invested in the creation of a museum exhibit to be installed at the Ann Arbor Hands-On Museum, a children's science center. Informal learning environments, like museums, align well with outreach efforts because they share many goals: to intrigue, educate, and inspire visitors. Science museums in particular have become more conscious of their role as an auxiliary to the education that occurs in traditional classrooms, striving to encourage interest in science[10] and to present science policy issues[9] that might not get addressed in the classroom.

This exhibit is divided into three parts, reflecting three interrelated fields required to bring a product to market: design, manufacturing, and marketing/business.[6,8] Each of the three parts of the exhibit was designed to emphasize the processes, tools, and careers of its field, through an interactive computer game as well as through hands-on exhibit components. The computer games were designed to be deployable both within the physical exhibit as well as online, to extend the outreach beyond the exhibit's physical location (see Figure 1).

The first game in the series, *Design Station*, invites the visitor to help employees of a design firm as they attempt to design a pen that will sell well. The visitor engages in market research and uses the results of that research to select pen components likely to please members of the market. The second game in the series, *Some Assembly Required*, introduces visitors to several characters engaged in different occupational roles related to manufacturing, such as a manufacturing engineer, a machine technician, and a floor manager. These characters introduce the visitor to the tools (e.g. injection molders and extruders) and tasks that he or she engages in – like machine calibration and the routing of materials to machines on a factory floor. The third game, *Business as Usual*, introduces the visitor to strategic planning, and asks the user to decide how to invest capital to best encourage sustained growth.

Figure 1. The *Design Station*, the first of three exhibits installed in the Ann Arbor Hands-On Museum.

The remainder of this section describes the design rationales underpinning the games in this exhibit, with particular attention towards the educational strategies used to support different types of learning objectives and the strategies used to make the game accessible to visitors of both genders. Section 2 describes our assessment goals for this project: how the collected data can be used to understand (1) **demographic** group performance, (2) **game-play** issues, and (3) **educational impact**. The engineering behind our data collection system is described in Section 3, analysis of the results is presented in Section 4, and Section 5 closes with overview of future research issues.

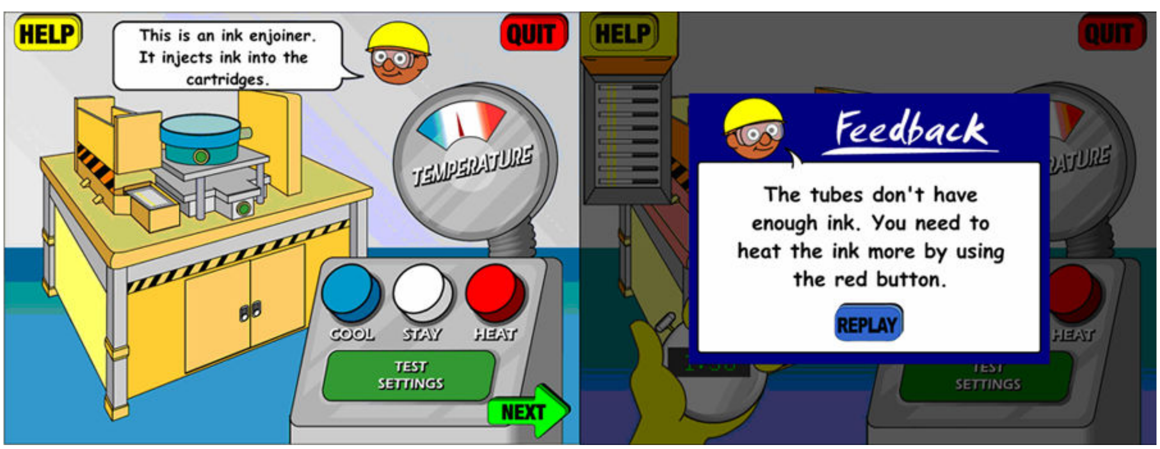

Figure 2. In *Some Assembly Required* Sanjay, the machine technician, introduces an ink enjoiner (left), and provides feedback during its calibration process (right).

1.1 Educational Game Design

Many researchers of learning in informal settings take a sociocultural stance that "emphasizes that meaning emerges in the interplay between individuals acting in social contexts and the mediators - including tools, talk, activity structures, signs, and symbol systems - that are employed in those contexts."[12] In our exhibits, the game itself serves as a mediator that presents a simulated social context and its associated "talk," activity structures, signs, and symbol systems. The visitor is asked to help a character in accomplishing a task, and in doing so is introduced to the terminology, processes, and values that employees in the character's profession consider important.[5] For example, the visitor may be asked to help the game character to calibrate the temperature of an ink enjoiner (see Figure 2) or to help with the routing of raw materials to the machines that process them on the factory floor (see Figure 3). We ask visitors to play roles, like that of a machine technician, in a context that emulates a real-world workplace so that they get a personal introduction not just to the content area, but to the sociocultural context the work takes place in.[7] This situated scenario approach is aligned with theories of learning in informal settings, wherein it is important for visitors to be able to make connections among personal, sociocultural, and physical contexts.[3]

Figure 3. Caitlin, the manufacturing plant manager, from *Some Assembly Required*.

The user engagement and learning is enforced by the reward structure embedded in the game – in order to score well, the user must gain an understanding of the required tasks (e.g. calibrating a machine to work efficiently, or designing the layout of a factory floor). In all of the games, the visitors receive more than just a numerical score: characters from the game provide them with detailed assessments as well (see Figures 2 and 4). These detailed assessments provide reinforcement and an opportunity for the visitor to reflect on his or her performance, which are key instructional strategies.[4] Another opportunity for reinforcement comes in the form of a "Bonus Quiz" at the end of each game. Three terminology-oriented multiple-choice questions are drawn randomly from a larger set of candidate questions and presented to the user. After he or she makes a choice, a brief explanation is supplied in response (see Figure 5).

The in-game scores for the bonus quiz and task execution serve a research purpose as well. Performance on the bonus quiz lets us know how well visitors have absorbed the terminology present in the exhibit. Likewise, we can use the game score as an indicator of how well the visitor has learned about a process found in the field. We can use it in this way because the score is tightly coupled to the execution of a task relevant to the career field, unlike other games that ask users to do something unrelated to the content area (like the flip-over card-matching games). For example, in *Design Station*, visitors are asked to go to a mall to conduct market research on people's preferences for pens, and to then use what they found to help select the components of a pen. A user who has internalized the task requirements will understand that it is important to talk to as many people as possible, and to select components that will please the largest number of people.

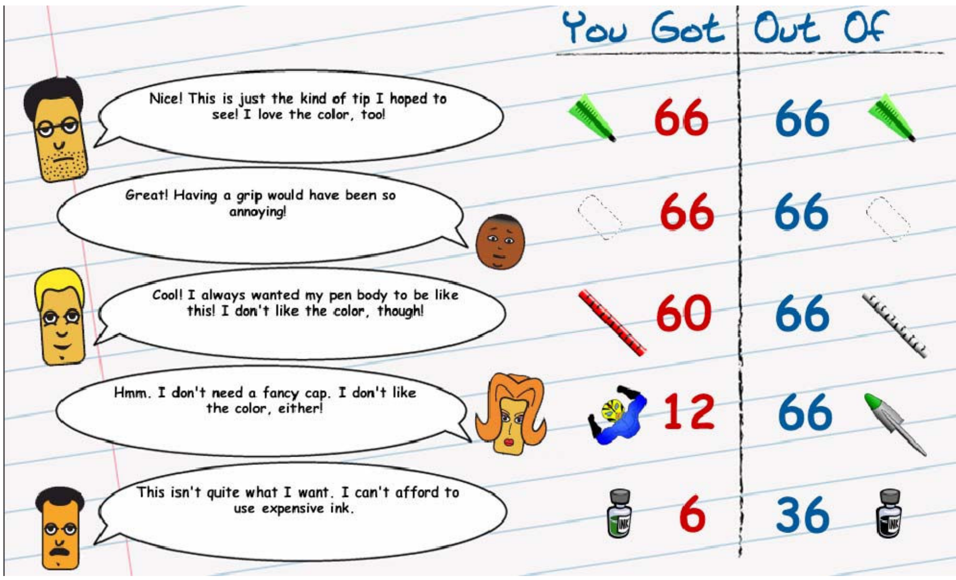

Figure 4. The score screen from *The Design Station*, showing how the user's pen component selections match up against the components desired by the most people.

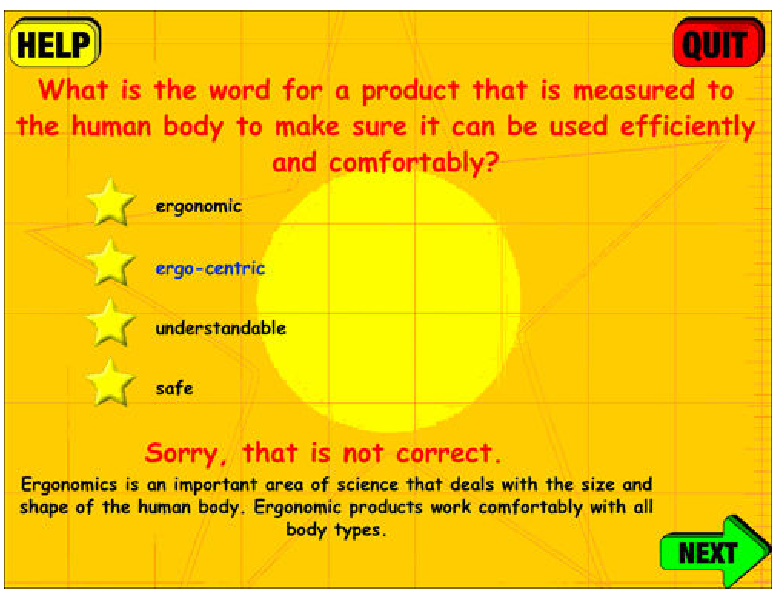

Figure 5. The feedback screen in the bonus quiz of *The Design Station*.

1.2 Game Design for Gender Equity

Special attention must be given towards designing the games to encourage equal participation of girls and boys, because there has been evidence that girls are less likely to become engaged by technology-heavy science center exhibits.[2] Just as children are able to recognize early on that certain toys are "intended" for a given gender, they are also able to recognize which computer games are "intended" for boys, and which are "intended" for girls. It seems that these judgments are largely made on the basis of the artistic and color schemes used.[13] In addition to initial impressions, girls may prefer certain types of game activities over others. There is some evidence that girls tend to prefer games that require cognitive skills that they are adept at, like matching, memory, and verbal skills,[13] and that they prefer games that center on creation rather than destruction.[1] Even the means provided for playing the game can introduce bias: girls seem to have more problems with certain input devices than boys. For example, although females perform just as well as males when using kinesthetic input devices, like touchscreens, they perform markedly worse with an abstracted input device like a mouse.[14]

We have made an effort to design our games to appeal equally well to both genders. To counter immediate impressions, we have been careful to stay away from using graphic styles that skew towards any obvious gender stereotypes, using neutral and primary colors and "cartoon"-style art. None of the tasks in our games are destructive, and often require visitors to rely on verbal and memory skills to better enfranchise female visitors. In the physical exhibit, we make use of a touchscreen interface instead of a mouse or trackball to make the game equally accessible, and the input elements are large, which obviates the need for precise pointing-device skills amongst our online visitors.

2.0 Assessment Data

Because this exhibit was created as an outreach effort, it is important for us to ensure that the games we create are accessible to our target age range, equally appealing to both genders, enjoyable to play, and are helping the visitors learn. For this reason, any assessment we do needs to be able to (1) ascertain the (self-reported) **demographics** of the online visitor population, (2) determine which portions of the game are problematic from a **game-play** perspective (and may require redesign to become more engaging), and (3) assess how successfully the target concepts were being communicated by the exhibit – i.e. its **educational impact**.

2.1. Demographic Data

In both the museum and online environments, it is impractical or impossible to truly establish the demographics of the visiting population. In the case of the physical exhibit, museum attendance figures can be used to get a picture of the proportion of adult and child attendees, but there is no way to know which of these visitors actually stop by the exhibit itself without carrying out an observational study. In the case of online visitors, there is no practical way to verify their demographic characteristics. For that reason, we must rely on self-reporting.

In either case, we must be sure that we are collecting visitor information ethically. In addition to seeking permission from the hosting institution, we need to be certain that we adhere to ethical standards for data collection over the internet.[11,15] According to these guidelines, it is best if no personally identifiable information, like names and addresses, are collected. For this reason we decided against having the users enter a name, even though it would have been useful in identifying returning visitors. The guidelines also recommend a clear disclosure of the fact that data will be collected, and an explanation of both the reason for the data collection and how that data will subsequently be used (for example, if the data will be made available to other researchers or institutions). For this reason, we made sure that as soon as an online user initiates a game, he or she is presented with a disclosure screen (see Figure 6). We attempted to make the language as easy to understand as possible, and because children may not legally be allowed to give informed consent, we make sure to mention that they need permission from their parents. We also allow users to opt out of the demographic collection, even though that may lead us to collect data on a self-selected population of users.

Figure 6. An image from the informed-consent screen of *The Design Station*.

2.2. Game Play Data

Much can be learned about the success of a game by watching users play it: this is why many professional video game developers spend a fair amount of money on user testing. With a relatively simplistic game, it is very easy to collect usage data, however. Our games all have a limited number of stages and choice points, and the interaction is via pointer clicks, so a record of all user activities can be confined to a small text file (less than 10 KB of space). By measuring the amount of time users spend in each stage of a game we can infer which parts of the game are liked or disliked, which portions of the game they are likely to quit from, and what choices they make when presented with alternate paths through a game.

2.3. Educational Impact

As mentioned in Section 1.1, if designed properly, the scoring mechanism built into the game itself can serve as an indicator of the degree to which a visitor has internalized a task. Likewise, bonus quiz score is an indication of how well the visitor has come to understand terminology presented within the exhibit. Because our desire to capture learning data had to be balanced against the desire to create an exhibit that would attract visitors, we opted not to include any pre-tests that might turn away visitors who were attracted to the notion of playing an interactive game. As a consequence, we are not truly measuring *learning* per se; rather, we are assessing the amount of knowledge that a visitor leaves the exhibit with. However, given how little the general populace knows about manufacturing and the relatively specialized exhibit content, these figures may be sufficient to make a case for the utility of the exhibit.

3.0 Data Collection

The collection of data from a computer game, even one implemented in a web-friendly multimedia development package, does not happen by default. Although web server logs can give statistics on traffic, followed links, and the like, current highly-interactive multimedia file types (like those produced with Macromedia Flash or Director) do not provide any data on a user's behavior. A web server will recognize when a multimedia file is requested (e.g. when the user clicks on a link to it), but all of the user's actions within that file go unreported – unless the developer engineers it to be otherwise. We used Macromedia Director to create both the exhibit and web versions of our games, and built in a mechanism to store the user's actions in the memory during gameplay. Essentially, every onscreen element was imbued with a "tattletale" behavior, so that the subject and time of each user click is reported on. We also record other game information (such as the final score and the questions answered by the user). When the user chooses to quit the game (or it times out from lack of activity), the user's data trace is stored more permanently. In the case of the exhibit, it is stored locally in a small text file.

In the case of the web version, the user's trace data is transmitted to a MySQL database via small PHP scripts.[8] To ensure that the PHP file requests actually reach the hosting server (as opposed to reaching a cached version of the file's output), each PHP file request is appended with a unique ID and a timestamp. In web version data collection has also a built-in "expiration date" so that the server is not indefinitely flooded with data.

4.0 Assessment Results

These results are based on the data we collected over a year from the physical installation of *The Design Station* at the Ann Arbor Hands-On Museum, from September 2004 until September 2005. During this period of time, the Ann Arbor Hands-On Museum had 148,190 visitors, and 16,983, or 11.46%, of them used *The Design Station*. Considering there are over 250 exhibits present in the museum competing for visitors' attention, attracting one out of every 10 visitors is notable.

4.1. Demographic Assessment

We were concerned about two things with respect to demographics: was the game accessible to all of our target age ranges, and was it equally accessible to both genders. By "accessible" we mean that user of all ages and genders are able to (1) complete the game (showing that they were able to play through without incident), and (2) score reasonably well on the game (showing that they are able to assimilate enough of the game's instructions to perform).

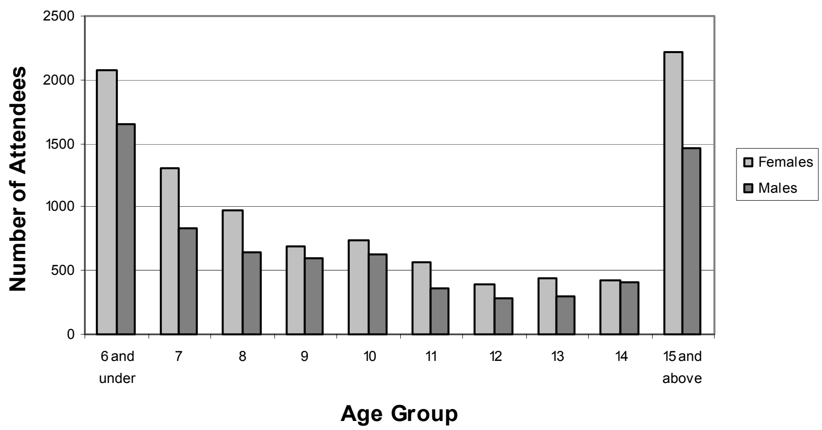

Figure 7. Demographic distribution of game players for the duration studied.

The distribution of visitors who made use of the game is depicted in Figure 7. Both numerically and proportionally, more children (13,302) than adults (3,681) made use of the exhibit. Physical observations made at the museum confirmed that nearly all of the 6-and-younger age group made use of the exhibit in conjunction with an adult, which accounts for why the 6-and-under results often parallel those of the 15-and-older age group. Overall, over half of the users who began playing completed the game (defined as those users who reached game stage 23, the score screen, or beyond). Figure 8 shows a breakdown by age group.

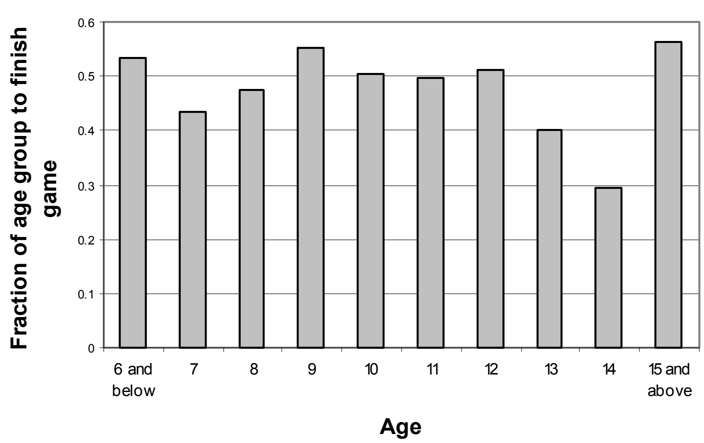

Figure 8. The fraction of each age group that played through the game.

Using this measure, it seems that the game appealed most to users from 9 to 12 years of age, adults, and adult-child dyads (nearly all young children play the game with the company of an adult). Examining the 13- and 14-year old users shows that the majority quit at stage 5 (see leftmost image in Figure 10), which is the end of the game instructions. The obvious conclusion is that the game, as described by the in-game avatar, does not sound appealing to older children. The relatively high rates of game completion for 7- and 8-year-olds (43.4% of 7-year-olds and 47.5% of 8-year-olds who begin the game, see Figure 8) implies that the game is accessible to even the younger end of the target age spectrum. Overall, slightly more females (51.6% of female users) than males (48.1% of males) completed the game, but virtually all of this difference can be explained by the fact that more adult females (56.8% of 15-and-above females) than males (44.3% of 15-and-above males) completed the game. For our target audience, there is no real difference, showing that the game is not more or less accessible to either gender.

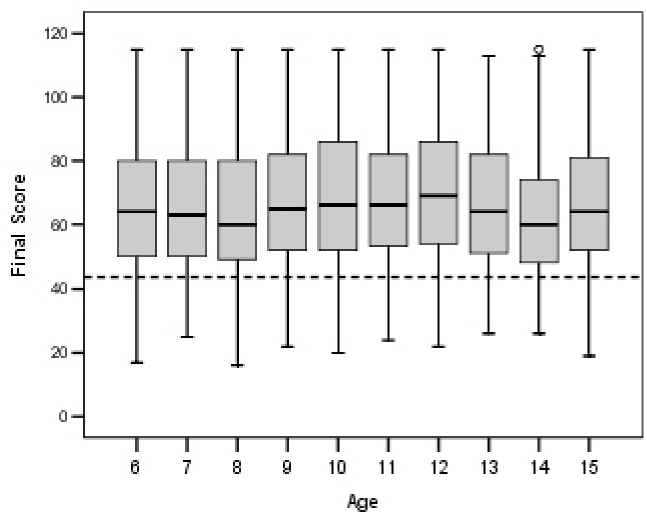

Figure 9. Average game score by age group. Dotted line indicates the expected score if the game were played randomly.

Even more encouragingly, the scores across the age groups were relatively uniform, further confirming that the in-game tasks were approachable by all age groups. (Recall that the game score is a reflection of the degree to which users understand the task requirements).The average game score obtained by all users was 66.1 (out of 120 possible points) and amazingly all age and gender groups performed very close to this figure (see Figures 9 and 12).

4.2. Game Design Assessment

As game designers we were concerned with making the game as enjoyable as possible to our visitors. Given the free-choice learning environment both museums and the internet provide, we were especially concerned that visitors would find the text-based explanation portions of the game boring (e.g. when task instructions or information on a profession are provided),. We were also concerned that the task activities in the game (such as the market research and the pen design portions) would not be appealing to certain visitors. One way we can determine this is by examining which stages of the game prompted visitors to quit early.

An analysis of the stages users quit at shows that for every age group aside from 14-year-olds, the very last stage in the game, 29, was where the majority of users quit (see Figure 10). This is a pleasing result, showing that the largest group of users preferred to play through to the end of the game. Stage 5, where most 14-year-olds quit, is the end of the instruction phase, just prior to the market research phase. Literally, the avatar is asking: "Are you ready to start?" on stage 5, so it is reasonable to assume that the users who quit at this stage did so because the activity that was just described in stages 3 and 4 does not sound appealing (see Figure 10). This is a satisfying result which counters concerns that fair number of users would quit during text-heavy expository and instructional stages; that is clearly not the case.

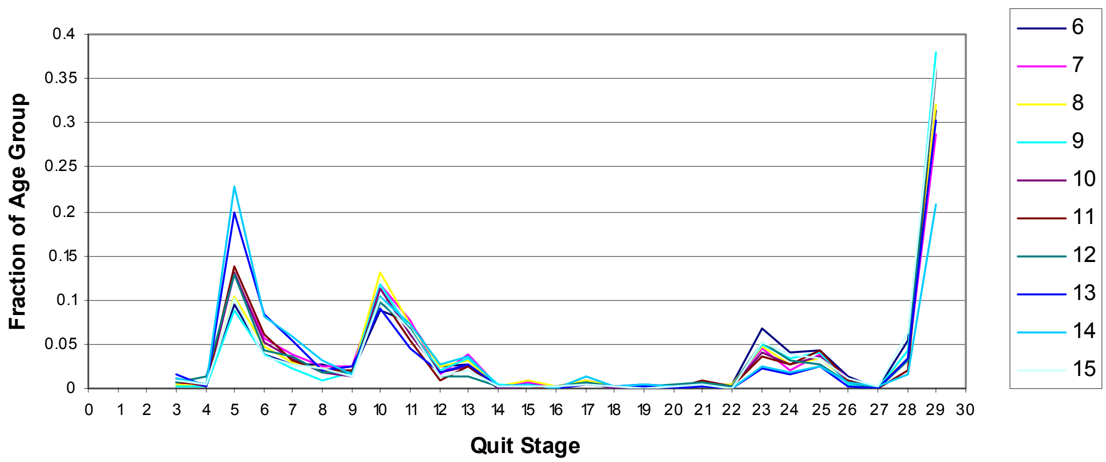

Figure 10. Frequency of game quitting for various age groups (stages 3 and 4 are the market research instructions).

4.3 Educational Impact Assessment

As seen in Figure 9, all user groups managed to perform very well at the tasks given them during the game, scoring above 60%. In this first game, those tasks were to a) conduct market research to discover people's preferences in pens and then b) apply what was learned in the selection of components for a pen to be manufactured. All groups performed much better than chance: were users to randomly select pen components, the expected score would have been 43.8% (represented by the dotted line in Figure 9 and shaded region in Figure 12). This shows that the users were able to understand the critical features of the tasks and execute them, thus gaining an understanding of the presented career field. From a game design perspective, this high performance across different age levels and different genders is exactly what we hoped to see: in designing this game we wanted the majority of the museum's attendees to find it accessible and enjoyable. The fact that no one age or gender group drastically out-performs any others, and that all groups perform well, argues that we attained our goal of creating a broadly-targeted educational game. Although the uniformity of the high performance across different age levels may seem unexpected, given the differing abilities of the age levels, the reason seems to be that the game offers three levels of difficulty, and proportionally more of the younger users chose the easy level, while proportionally more of the older user chose the hard level (see Figure 11). This self-selection seems to have balanced out the performance amongst age groups.

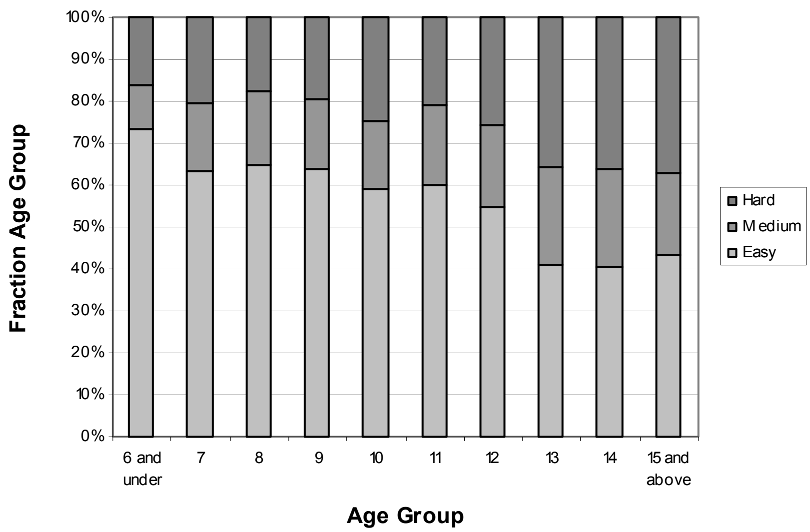

Figure 11. Breakdown of the game difficulty levels chosen within each of the age groups (notable preference of older players for higher challenges).

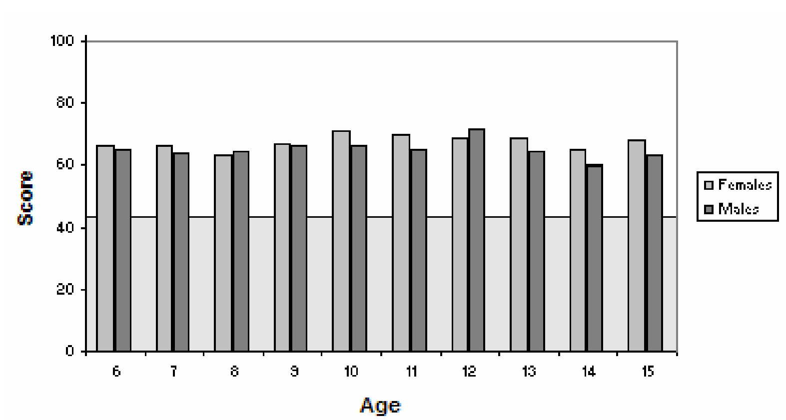

Figure 12. Performance of various demographic groups relative to chance (shaded region of the plot area corresponds to scores from random game plays).

With respect to knowledge of terminology, the performance of those users who completed the "Bonus Quiz" is high: the average is 2.44 correct out of three quiz questions. (The quiz is comprised of three questions reflecting terminology or other factual knowledge presented in the exhibit. The three questions are selected randomly from a pool of 15 questions, so repeat players will not get the exact same three questions). This high performance should not be surprising, however: the "Bonus Quiz" is another realm where self-selection plays a role. A large majority of users, 79.1%, opted not to take the quiz. This suggests that we should use a better incentive than an increase in the “final” score (the sum of the game score and bonus points, which is used to determine the high scores posted at the end of the game).

5.0 Discussion and Future Work

It seems that, unlike traditional assessment techniques, the use of built-in data collection necessitates that the designers weigh the visitor’s experience with their software against their data collection options. For example, an optimal feature from a data collection standpoint would be to incorporate a pre-test “quiz” into the beginning of the software so we could compare pre- and post-“quiz” performance. The sheer number of users who opted *not* to engage in the post-“quiz” factual assessment, however – nearly 80%, suggests that the inclusion of a pre-test will definitely turn some users away. It is necessary to balance a need for useful data against an assurance of an enjoyable user experience – in general, the more demanding the task, the more useful information researchers can get out of it, but the likely consequence is diminished user enjoyment.

There may be a way to find a compromise between the needs of the researchers and the needs of the users, however. One advantage to the development framework our games use is that regular, external text files can be used to configure options within the game itself (when the game begins, it consults these text files to import variables for use during run-time). We can use these files to “shut off” certain data collection features, like a pre-test “quiz,” after a significant body of data has been collected, thus allowing researchers to acquire the data they need while inconveniencing a smaller number of users.

Determining user preferences is another area where text-based configuration files can help: when combined with our automatic data collection, the use of these files can be quite powerful for resolving open questions about certain game design features. For example, we could make use of a lengthier expository sequence 50% of the time the game is played, allowing us to see if this new element causes more users to quit the game early. If we determine that the sequence does have a negative impact, we wouldn’t need to recompile the game to remove it; we could shut off the sequence by altering just the contents of the text file. This approach can help answer a great many open game design questions while simultaneously collecting data for research purposes, and will be incorporated into the other games currently under development. Effectively, such features can also enable carrying out experimental studies which would allow to shed light on impact of particular elements of contents within the games and better gauge user experiences.

We hope that this work has introduced the challenges, solutions, and open questions involved in assessing the behavior of visitors to a highly-interactive computer-based exhibit. In an era where students display a marked lack of enthusiasm for science and engineering topics and careers, it may be strategic to target venues outside classrooms for outreach efforts. Informal learning contexts like museums allow visitors to follow their curiosity and to engage in role-playing, behaviors not ordinarily seen in classrooms. The accompanying challenge, though, is that with an absence of traditional assessment tools (like homework and tests) it is that much harder to determine if the outreach effort is successful. The intersection of informal learning environments, highly-interactive computer games, and automated data collection gives us an opportunity to understand how to tune outreach efforts to function well in these venues.

References

1. American Association of University Women (AAUW). *Tech-Savvy: Educating Girls in the new computer age,* Washington, D.C.: AAUW, 2000.

2. Carlisle, R. W. What Do School Children Do at a Science Center?, *Curator 28*, 1, 1985.

3. Falk, J. H., & Dierking, L. D. *Learning from Museums: visitor experiences and the making of meaning*. Walnut Creek, CA: AltaMira Press, 2000.

4. Gagne, R., Briggs, L. & Wager, W. *Principles of Instructional Design (4th Ed.)*, Fort Worth, TX: HBJ College Publishers, 1992.

5. Gee, J. P. *What Video Games Have To Teach Us About Learning and Literacy,* New York: Palgrave Macmillan, 2003.

6. Kaynar, I., Pasek, Z., and Lyons, L. "Creating an Informal Engineering Education Experience: Interactive Manufacturing Exhibit," *Proceedings of International Conference on Engineering Education (ICEE2004)*, Gainesville, University of Florida, October 17-21 2004.

7. Lave, J., & Wenger, E. *Situated Learning: Legitimate Peripheral Participation,* Cambridge, UK: Cambridge University Press, 1990.

8. Lyons, L., and Pasek, Z. "Gauging Visitor Behavior at an Interactive Engineering Exhibit,“ *Proceedings of the 2006 Conference of the American Society for Engineering Education (ASEE2006)*, Chicago, IL, June 18-21, 2006.

9. Macdonald, S. & Silverstone, R.. "Science on Display: The Representation of Scientific Controversy in Museum Exhibitions." *Public Understanding of Science 1*, 1, 1992.

10. Paris, S., Yambor, K, & Packard, B. "Hands-On Biology: A Museum-School-University Partnership for Enhancing Students' Interest and Learning in Science." *The Elementary School Journal 98*, 3, 1998.

11. Sandvig, C. & Murase E. Social Research Through the Unobtrusive Observation of Network Traffic: Methodological and Ethical Challenges. In *Proc. AIR 2000*, AIR 2000. Consulted January 27, 2005. http://aoir.org/members/papers/Sandvig_Murase--AOIR_Paper_2000.pdf

12. Schauble, L., Leinhardt, G., & Martin, L. "A Framework for Organizing a Cumulative Research Agenda in Informal Learning Contexts." *Journal of Museum Education 22*, 2&3, 1997.

13. Sherry, J., Holmstrom, A., Binns, R., Greenberg, B., & Lachlan, K. Gender Differences in Video Game Use and Preferences. In *Proc NCA 2003*, NCA, 2003.

14. Tan, D., Stefanucci, J., Proffitt, D., Pausch, R. Kinesthetic Cues Aid Spatial Memory. In *Proc. of CHI 2002*, ACM, 2002.

15. Thomas, J. “Introduction: A Debate about the Ethics of Fair Practices for Collecting Social Science Data in Cyberspace.” *The Information Society 12*, 2, 1996.