Making Player Engagement Visible: A Multimodal Strategy for Game Experience Research

Jeffrey Bardzell* | Shaowen Bardzell* | Tyler Pace* | Jeremi Karnell°

*Indiana University School of Informatics
1900 E. Tenth Street, Eigenmann 937
Bloomington, IN 47406
{ jbardzel | selu | tympace } @indiana.edu

°One to One Interactive
529 Main Street
Charlestown, MA 02129
JKarnell@onetooneinteractive.com

ABSTRACT

This paper describes an industry-academic collaborative research initiative, focused on determining useful measures of "engagement" and experience with interactive digital media (e.g., video games, virtual worlds, and digital devices, etc.). Using newly designed hardware and software, the research initiative addresses the relationships among neurological, physiological and cognitive assessments of engagement in ongoing and short duration gaming experiences. It is a centerpiece of an iterative strategy toward understanding and modeling relationships among different engagement measures. The research will lead to design proposals for model-based assessments of engagement calibrated to individuals' responses.

Categories and Subject Descriptors

H.5.2 User Interfaces: Evaluation/methodology. H.5.m. Information interfaces and presentation: Miscellaneous.

General Terms

Measurement, Experimentation, Human Factors.

Keywords

physiological, neurological, games, Quantemo, affect, emotion, usability, user experience, engagement

1. INTRODUCTION

HCI as a discipline has tended to focus on designing to increase efficiency and reduce frustration at the single user, single interface, or task and session level. However, the advent of interactive game experiences has led to a call for new methods. It has also changed notions of "use" and the "user." Video games are not used instrumentally to accomplish another task, as use of a word processor is instrumental to the composition of a legal brief; instead, its "use" is the experience it affords, and its ostensible content is often almost incidental.

How has the HCI world responded? There has been a call to move from thinking about activity to thinking about experience [10], and then to emotion [12], fun [2], engagement [5] and to enchantment [11]. These elusive concepts are hard to measure, which poses problems for traditional efficiency-oriented design and usability evaluation methods.

Reflecting philosophical and paradigmatic differences in reasoning about human experience, engagement and experience measures vary from qualitative, cognitive assessments elicited from people in surveys, interviews and diary studies, to observations of posture, body orientation and activity, to physiological measures (Galvanic Skin Response (GSR), Respiratory Rate (Br/M), Electrocardiogram (EKG)/Cardiac Measures, Electroculogram (EOG)/Eye-Tracking, Electro-myogram (EMG)/Facial Recognition) and neurophysiological measures (Electroencephalogram (EEG)/Brain Activity). It is clear a mixed-methods approach is best for addressing human engagement and experience, and that methods need to be balanced differently depending on the nature of the activity, the context of the activity, and the time frame for assessment.

The evaluation of user experience and engagement in gaming environments is a diverse and growing research area in HCI [1, 6, 13, 16]. Researchers, developers, producers and marketers are all interested in developing critical and productive methods for evaluating and assessing user experience in games. Collaboration and resource sharing between these groups is necessary for the fruitful future of game experience evaluation. This paper presents one such collaboration and a novel user experience methodology developing through the partnership.

Indiana University's School of Informatics and One to One Interactive, a digital media marketing and research firm, began the partnership in 2007 to pursue a new line of research into user experience and emotional engagement with digital devices, video games, and virtual worlds. The partnership brings together academic expertise in subjective experience evaluation with industry leading technologies for physiological and neurological measurements of user engagement.

2. USER EXPERIENCE EVALUATION METHODS

User experience evaluation methods in games can be divided into two broad categories: objective and subjective. Subjective measures record a user's perception, generally self reported, of the media at hand. Objective measures are not dependent on a user's perception and generally include independent measures such as the passage of time or number of mouse clicks to complete a task [7].

2.1 Objective Measures

Objective measurements of user experience in games come largely in the form of traditional user evaluation or usability testing and the more recent use of physiological and neurological measurements of game players.

Objective user evaluation methods include eye tracking, task completion timing, click counting, cursor monitoring, screen capture, video capture, etc. Traditional user testing, represented

through a combination of the previously mentioned techniques, are established practices with the game development and research communities and will not be discussed in detail in this paper. Alternatively, this paper focuses on neurological and physiological measurement techniques which are seemingly less common in both professional and academic practice for evaluating player experience of games.

Neurological measures of user experience focuses on brainwave activity as detected through electroencephalography (EEG). Brainwave activity is broken down into six categories: delta, theta, alpha, low beta, midrange beta and high beta. Each type of brainwave is associated with a particular state of physical and emotional arousal [15].

The most common physiological measures of user experience are galvanic skin response (GSR) and heart rate (HR) monitoring. GSR, also known as electrodermal response (EDR), psychogalvanic reflex (PGR), or skin conductance response (SCR), is a method of measuring the electrical resistance of the skin. There is a relationship between sympathetic activity and emotional arousal. GSR is highly sensitive to emotions in some people. Fear, anger, startle response, orienting response and sexual feelings are all among the emotions which may produce similar GSR responses [19].

Heart rate monitoring is a simple measure of the number of heart beats per minute. HR has a well established history in media research due to the numerous inferences that can be made from cardiac activity. Patterns in HR demonstrate attention, effort, arousal and emotion. Additionally, HR is an indicator of short-term attentional selection and long-term attentional effort [14].

Neurological and physiological measures of user experience with media offer numerous benefits over purely subjective measures and traditional usability testing. Physiological and neurological measures may reveal contradictory information to that of self-report data. Physical measurements may reveal changes too subtle to be made available to the participant's conscious awareness. Additionally, physical measurements are constant throughout a user session whereas self-report data is usually only collected at the end of the use experience. Finally, physical measurements do not depend on language, interfere with message processing or require memory [14].

Disadvantages of physiological and neurological measurements well documented in HCI literature. First, the measurements are inconsistent and prone to variations between subjects. Second, physical measurements result in large data sets where determining significant events posts a methodological challenge. Finally, interpreting significant measurements is difficult. Different mental states produce similar physical readings making the objective determination of a precise mental state a challenging task [14].

2.2 Subjective Measures

Subjective user experience evaluation methods include surveys, interviews, contextual inquiry and other ethnomethodological techniques [9]. Additionally, methods of measuring affects and user engagement are also explored in the domain of affective computing [4, 8]. By distinguishing two models of affect (i.e., information vs. interaction), Boehner et al. reiterate the need for affect and engagement research in HCI to take into account users' awareness and subjective interpretation of emotional response to technology [2].

Generally speaking, subjective measures offer insight into the user experience directly from the user and allow users to express the complexity and depth of their interactions with media in their own terms. However, subjective measures are usually performed in retrospect and may not be reflective of the entire media experience, but rather only with the best and worst memories of the experience. Additionally, self-report information may be influenced by the pressure to respond positively or with political correctness [14].

3. OUR APPROACH

Given that emotion and engagement are both biological and socially and culturally constructed and enacted, our approach to player experience evaluation of digital games and media is a triangulation strategy of neurological and physiological measurements combined with traditional user testing and subjective, reported feedback techniques to identify patterns between the information and interaction models. Neurological and physiological measures are managed by the Quantemo™ system developed and maintained by One to One Interactive.

Quantemo's MindKit™ platform is used to extract, filter and amplify brainwave (EEG) signals and convert that information into digital emotional state outputs. MindKit™ can be used for monitoring and analysis of human brainwaves and mental training. The EEG signals read by MindKit™ detect on the forehead via points Fp1 or Fp2 of the electrode placement system developed by the International Federation in Encephalography and Clinical Neurophysiology.

Quantemo's LifeShirt™ system is an ambulatory, multi-sensor, continuous monitoring system for collecting, analyzing and reporting physiological data. The LifeShirt system is built upon Respiratory inductive plethysmography (RIP) measurements to capture a highly accurate view of the subject's breathing in both clinical and non-clinical settings. Additionally, the LifeShirt™ system is able to collect reliable objective physiologic data through various sensors, including function, tidal volume, respiratory rate, electrical activity of the myocardium via a 3-lead electrocardiogram (ECG), and participant activity/posture via a tri-axial accelerometer [18].

In addition, we will also use traditional evaluation methods (e.g., eye tracking, click tracking, and digital video capture technologies) to monitor the use behaviors of research subjects as they interact with games and virtual worlds. These methods are then combined with traditional quantitative/qualitative methods (e.g., surveys and interviews, etc.) as well as novel ways for users to articulate and interpret their emotions through tangible interactions to correlate observed behavior with self-report information [2].

4. RESEARCH PLAN

Beginning November 2007 and over the next year, we will be working to address two problems. First, we will be using these techniques as described to study player engagement with games and game-related phenomena (e.g., machinima). At the same time, and more importantly, we will be working to improve our methodology, learning how to manage, correlate, and interpret the various modalities available to us: interviews,, click tracking, eye movement, and the physiological and neurological data as well.

While in theory we could focus on one or the other first, our goal is to complete a half-dozen to a dozen short studies where we deploy these research techniques to study specific questions relating to player engagement with games. Through the sequence of these studies, we will iteratively improve our research approach. That is, the study of specific, concrete issues in player engagement and our techniques for correlation and interpretation of data will be mutually informing.

Our immediate research agenda for the winter of 2007/08 includes the following:

- Comparison of user engagement between user/amateur created games and professional "triple A" games. Amateur games utilize different design approaches, tools, budgets and testing techniques than large professional games. Do amateur and professional games engage users in different ways?
- Identify "flow" states [3] during game tutorials. Video game tutorial are often overlooked by game players and game developers, despite their importance as educational and orientation tools for the player. Identifying successful tutorials and their flow states may lead to a better understanding of how to design more engaging tutorials in the future and yield insights into making games more broadly accessible and appealing.
- Determine responsiveness to in-game advertising in casual, serious and professional games. In-game advertising is becoming a serious revenue model for games of all types, but little is understood on the effects of in-game advertising on player experience and engagement with the game. Alternatively, the marketing effects of in-game advertising on players are equally under-researched.
- Measure reactions to fan media for fans and non-fans. Fans react in obvious, pleasurable ways towards their favorite media but the effects of fan media on non-fans is not well understood. Understanding how non-fans react to fan media may allow better development of media that can appeal to non-fans and convert them to fans.
- Assess the engagement qualities of casual games across age and gender. Casual gaming is much more prevalent among women and older individuals than traditional or "hardcore" games. Determining the types of player experience of casual gamers with casual and traditional games will mutually inform the future development of each genre.

5. FUTURE RESEARCH CHALLENGES

Alongside the goal of our using the Quantemo™ system to evaluate player experience with games is, of course, the desire to make this kind of research replicable and usable by other teams in other contexts. A significant part of that initiative is the development of visualization techniques that will help researchers perceive relationships among the different modalities. As we run the experiments and analyze the data, we are also participating in the design of visualizations.

The result of all of this work will be the development of a series of protocols supported by a software application for researchers. This application will provide access to the data, offer various 2D and 3D animated visualizations, and include other data discovery tools that will facilitate analysis and interpretation of the complex and hard-to-see phenomenon that is player engagement.

6. REFERENCES

[1] Bernhaupt, R., Eckschlager, M., & Tscheligi, M. Methods for evaluating games: How to measure usability and user experience in games? *Proc. of the International Conference on Advances in Computer Entertainment Technology ACE'07*.

[2] Boehner, K., DePaula, R., Dourish, P. & Sengers, P. Affect: From Information to Interaction. AARHUS'05.

[3] Csikszentmihali, M. *Flow: The Psychology of Optimal Experience*. (1990). Harper Perennial.

[4] Desmet, P. Measuirng emotions. In Blythe, M., Overbeeke, K., Monk, A., and Wright, P. (eds). *Funology: From Usability to Enjoyment*. (2004), 111-124. Kluwer Academic Publishers.

[5] Douglas, Y., & Hargadon. The pleasure principle: Immersion, engagement, flow. *Proc. of the eleventh ACM on Hypertext and Hypermedia*. HYPERTEXT'00. (2000), 153-160.

[6] Gilleade, K., & Dix, A. Using frustration in the design of adaptive videogames. *Proc. of the 2004 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology*. ACE'04.

[7] Hornbæk, K. 2005. Current practice in measuring usability: Challenges to usability studies and research. International Journal of Human-Computer Studies 64, (2007), 79-102.

[8] Isbister, K., & Höök, K. Evaluating affective interfaces: Innovative approaches. CHI'05.

[9] Kuniavsky, M. Observing the User Experience: A Practitioner's Guide to User Research. Morgan Kauffman, 2003.

[10] McCarthy, J. & Wright, P. *Technology as Experience*. The MIT Press, 2004.

[11] McCarthy, J., Wright, P., Wallace, J., & Dearden, A. The experience of enchantment in human-computer interaction. *Personal and Ubiquitous Computing*. v.10(6). (2006), 269-278.

[12] Norman, D. *Emotional Design: Why We Love (or Hate) Everyday Things*. Basic Books, New York, NY, 2004.

[13] Pagulayan, R., Steury, K., Fulton, B., and Rmero, R. Designing for fun: User-testing case studies. In Blythe, M., Overbeeke, K., Monk, A., and Wright, P. (eds). *Funology: From Usability to Enjoyment*. (2004), 137-150. Kluwer Academic Publishers.

[14] Ravaja, N. Contributions of Psychophysiology to Media Research: Review and Recommendations. Media Psychology 6, (2004), 193-235.

[15] Teplan, M. Fundamentals of EEG Measurement. Measurement Science Review 2, (2002), 1-11.

[16] Thompson, C. The science of play. In *Wired*. September 2007.

[17] Ward, R.D. and Marsden, P.H. Physiological responses to different WEB page designs. International Journal of Human-Computer Studies 59, (2003), 199-212.

[18] Wilhelm, F., Roth, W. and Sackner, M. The LifeShirt: An Advanced System for Ambulatory Measurement of Respiratory and Cardiac Function. Behavior Modification 27, (2003), 671-691.

[19] Wilson, G. and Sasse, A. Straight from the Heart: Using Physiological Measurements in the Evaluation of Media Quality. Proceedings of the Society for the Study of Artificial Intelligence and the Simulation of Behaviour (AISB) Convention 2001, Symposium on Emotion, Cognition and Affective Computing, March 21st - 24th 2001, York, UK, 63-73.