Crafting User Experience via Game Metrics Analysis

Anders Tychsen
Center for Computer Games Research
IT University of Copenhagen
anty@itu.dk

ABSTRACT
Game metrics are numerical data obtained from the user-interaction with the game software, and form a valuable data source to user experience research and –design, because these data offer quantitative, time-stamped information about the specific behavior of players of computer games. By combining this game metrical data with traditional user experience measures – biofeedback, surveys and usability-methods – it is possible to directly link game experience with design elements.

Author Keywords
Metrics, user experience, instrumentation, game design

ACM Classification Keywords
K.8 [Personal computing]: Games; J.4 [Social and Behavioral Sciences]: Psychology.

1. INTRODUCTION

Digital games are complex pieces of software that give rise to a diversity of potential user experiences. This means that it is challenging to know precisely what the experience of a given user will be when playing a computer game. As there are no generally accepted models for how games create experience, nor for how specific elements of games, or specific types of games, are related to specific components of game experience, game designers are therefore faced with a demanding challenge when developing games.

In the ideal world, game designers had access to libraries of knowledge based on empirical experiments with different types of game designs, and recordings of how the specific details in the designs promote specific components of the user experience, within a variety of user profiles. Lacking such extensive libraries, it becomes a key interest of game development to be able to test user experience of games under production: By linking specific features of the game under development with the user experience, it is at least possible to evaluate whether a given feature works or not, even if user testing cannot always say why this is the case.

Game testing during- and post-production has been performed for decades, however traditionally using informal methods. Within the last decade or so, a variety of methodologies have however been adapted from HCI to assist with this process, for example different forms of usability-testing [17]. These methods, e.g. audiovisual recording or interviewing players about their game playing experience and how the game design impacted on it, come with a large set of limitations. Audiovisual recording is time-consuming to analyze, with everything having to be done by hand, and is limited in that not every action of the player in the game world can be tagged and measured. Post-game interviews or surveys suffer from the problem that they are difficult to relate to specific of design. After a 50 minute play session a player may be asked what they did during the session, but their recall of the events of the game will be to greater or lesser degree imprecise, and their memories already biased. Using smaller in-game surveys in conjunction with limited playtime intervals may alleviate some of these problems, however, as yet no golden rule has emerged as to what playtime intervals that should be used.

Similarly, the past five-ten years has seen a substantial amount of debate about the nature of **user experience** in games (game experience), the elements it is comprised of, and various suggestions for how to measure it [11]. A lot of these are survey-based or rest on interviews and/or audiovisual recordings of the game sessions. The usefulness of these measures in an industrial design-process is a subject of ongoing debate.

Biofeedback measures offer another way of measuring user experience, through neurological and physiological reactions taking place in the player, e.g. [16]. However, measuring Galvanic Skin Reduction (GSR), EEGs etc., cannot provide as nuanced feedback as qualitative interviews or surveys, which enable researchers to dig into the underlying causes of players' reactions. Combining physiological indexes with survey data holds potential to develop models (function approximator) of specific user experiences, e.g. [16,22]

In summary, not only is the composition and causes of game experience not agreed upon, and few empirically grounded measures of game experience exist; it is also challenging to obtain accurate and clear data - in an unobtrusive manner - about the actual behavior of players within the virtual worlds of digital games. In essence, it is difficult to obtain knowledge about what players did when playing the game, and how meeting different game design elements affected their experience of interacting with the game.

Methods such as those discussed above are limited in that test managers can only hand-code so much information, and analysis of e.g. screen capture to a high level of detail is incredibly time consuming and not a good solution for the quick and effective game testing process required in the industry [17,19]. A potential solution is however presented in the automated collection and analysis of game metrical data, i.e. instrumentation data about the user-game interaction: **Game metrics.** These are numerical data obtained from the user-interaction with the game software itself, and can be recorded at different degrees of temporal and spatial resolution and even aggregated [8,21].

Metrics provide objective data on the interaction between players and games, and potentially any action the player takes while playing can be measured, including which buttons that are pressed, the movement of player-characters within the game world, or which weapons that are used to eliminate opponents. Within the past few years, this form of data has gained increasing attention as a means for obtaining detailed records of the player-game interaction. Metrics supplement existing methods of games-based user research, e.g. usability testing (measuring ease of operation of the game) and playability testing (exploring if players have a good experience playing the game) by offering insights into how people are actually playing the games under examination.

What game metrics offer is the opportunity to tie game experience to specific game design features. For example, if GSR recordings indicate high levels of arousal during three 15-second intervals in a 10-minute game session, metrics tracking software can show exactly where the player was positioned during these intervals, and what they were doing – action by action. Designers can then utilize this knowledge to develop heuristics for creating situations with high levels of arousal (an expert-knowledge approach). This is a primitive example – repeated uses of the exact same situation will likely cause diminishing returns in terms of arousal rate in players, as is the case with the surprise-attack-in-the-dark-routine of *Doom 3*, presumably the levels of arousal would decrease as players recognize the pattern.

Within AI research recent investigations have centered on creating automated systems adapting gameplay to the specific user, e.g. [23]. Resulting are as yet limited to simple non-commercial gameplay forms, however, combining these approaches with extensive libraries of metrics-biofeedback linked data

The purpose of this short paper is to introduce the concept of game experience, as well game metrics and how they are recorded and analyzed, to general HCI-practitioners, and to briefly reviews the current state-of-the-art. Furthermore, to introduce the idea of combining measures of user experience with game metrics to inform game design. The ideas presented stem from a recently established collaboration between the *Center for Computer Games Research, IT University of Copenhagen*, and game developer *IO Interactive*, who developed e.g. *Hitman* and *Freedom Fighters,* focusing on integrating metrics in industrial and academic game testing and user research.

2. GAME EXPERIENCE

Game experience is – not unlike user experience - a nebulous concept, which has garnered massive discussion and interest from theoretical and empirical angles alike. One of the most debated means of providing a measure for the overall gaming experience that crosses game forms and formats is "fun" [1,14]. It is intuitively obvious that games should be fun to play, as people play games because they enjoy them [18]. However, this is a problematic viewpoint and it has insofar not been possible to create operational procedures for game analysis, testing and design based on the concept of fun (the possible exception lies with approaches such as Entertainment Modeling, however so far this work has only considered simple arcade or augmented-reality games, not commercial-grade titles [22]).

First of all, the concept of fun is multi-modal, consisting of a lot of different motivations and reception factors that vary – at least - as a function of the game, the player and the play context. The result is that developers in the games industry tend to utilize their own experiences and those of their colleagues when designing games to be fun, rather than an objective and grounded approach.

Secondly, Dibbel [6] emphasizes the problematic nature of assuming that fun equals the gaming experience. Apart from being a difficult concept to harness in relation to evaluating gaming experiences, games offer other experiences that can be just as attractive: Challenge/mastery, teaching, bringing people together (social connection) or experiencing something unusual – even frustration. Fun has been viewed as the overall, assembling factor of the gaming experience; however, an alternative viewpoint argues that fun is just one of many potential types of experiences that can be derived from game playing [12].

A number of other concepts have been proposed as the umbrella that qualifies as an overall measure of gaming experience. For example, **enjoyment** as a concept within media studies covers media experiences broadly. The academic study of enjoyment (including emotional reactions) is a key field in communication theory, and notably motivational psychology contains substantial amounts of work that is useful in games, which however remains comparatively limited in its implementation in games testing and –evaluation. The few published works on the subject focus on modeling fun based on psychological theories/principles, e.g. [12,22,23].

Entertainment is conceptualized as an affective response to entertainment products, and also as a human activity that can be: *"Influenced, triggered, and maybe even shaped by the media product that is selected"* [2]. As a human activity, it contains physiological, cognitive, affective and behavioral components. Sweetser and Wyeth [20] argued that the plethora of models suggested within media theory to explain enjoyment, e.g. transportation theory, do not provide sufficiently well-rounded models of enjoyment. Instead, the authors highlighted the need to consider game experience in a broad context, and argued that enjoyment is the ultimate goal of games, adapting **flow theory** [4] to create a qualitative model for measuring flow in gameplay. The argument that flow equaled enjoyment was however criticized by [10], who noted that: *"Game enjoyment represents a broader set of experiences besides flow"*. Cowley [3] further criticized Sweetser & Wyeth [20] missing out some of the important aspects of flow at the elementary level. Cowley [3] further noted that the existing models that compare game elements with the eight dimensions of flow are not rigorous: Mapping between systems should occur at the level of systems.

Another problem with the concepts of fun, entertainment and flow is that neither of these intuitively fit with the experience of **negative affect** (**negative valence** when considering biophysical measures [16]) during game playing, e.g. aggression, frustration and stress [e.g. 8]. It could be argued that these are not necessarily experience components that game design should aim for, although this is not a simple issue, e.g. in relation to aggression. However, the problem stands that these components of the game experience exist. It therefore does not make sense to discuss game experience as a whole without considering these potential components.

Because of the variety of factors that the gaming experience can contain; the variability of the situation or context the gaming activity takes place in, the players, and the huge span of gameplay variations within games themselves, it remains unresolved whether it is possible to create a unified model of the gaming experience that covers all forms and formats of games, nor for testing its quality (although attempts have been made e.g. [11].

3. BIOFEEDBACK

Game experience is a psychological and physiological process. When playing a game, specific neurological and physiological reactions take place in the player. When discussing game experience components such as emotion, what is being referred to are biological responses in the game players brain and body, and some of these can be measured either directly or indirectly. For example, the emotional response to a digital game can be measured using **physiological** measures such as skin temperature and galvanic skin reduction [e.g. 16]. **Behavioral** responses [10]. **Neurological** responses can be captured using e.g. electroencephalograms (EEGs), functional magnetic resonance imaging (fMRI) [7]. The application to measuring aspects of gaming experience such as challenge, fun, boredom and engagement is relatively new, however, these measures have the inherent strength that they can present objective data on the user experience, thus eliminating one of the core problems with traditional HCI-methods such as usability. However, biofeedback measures cannot provide the nuanced feedback possible with qualitative methods, which enable researchers to dig into the underlying causes of players' reactions and thus complement biofeedback methods.

4. GAME METRICS

Game metrics is a term that is directly associated with metrics in general productivity software testing [15], however in the specific context of game development and –testing. Game metrics can therefore be related to user interaction with the game software, as is the case here, but also mean data of engine performance, sales or project progress. The former type of game metrics are basically User-Initiated Events (UIEs) [5,13], for example movement within the virtual environment, use of the game interface, use of specific in-game abilities, interaction with entities (agents) or objects in the game world or other players, etc. It should be noted that all games constrain the freedom of the user to interact with the game content in some ways, notably via the game mechanics. Metrics are objective; can be collected in large numbers and map to specific points in a game. In comparison, player-based feedback has much less resolution and is inherently biased due to individual preferences. Game metrics contain however an inherent limit in that these data cannot provide any contextual data. For example, whether the player is having fun or not, is male or female or whether another player is watching the screen – a metrics tracking tool can only record information from the specific game software. In essence, metrics cannot account for **social factors**, **the state of the user**, and of course any components of **game experience** (**emotional, sensorial, cognitive or behavioral**). Therefore, metrics need to be combined with other data sources in order to provide linkage between the user behavior – and thus game design elements – and game experience. This in a similar way that biofeedback data has been combined with survey data [22,23].

User-interaction metrics can take different forms. Some can be recorded on a continual basis, e.g. movement in the virtual world, or be recorded using specific frequencies, such as the location of the player avatar/-s (virtual representation/-s of the player) every 3 seconds. Metrics can also register triggered events, e.g.

every time an avatar jumps or shoots a weapon. Irrespective of the specific metric, a range of additional data are useful to capture, notably a time stamp, the spatial coordinates of the avatar at the time a metric recording is triggered, the originator of the metric (for example which player that fired the gun), the camera angle and vector of player avatar movement, and the content of the recording event (i.e. the metric recorded). The choice of which metrics to record and whether to record them as event sets or as aggregate data sets, depend on the game in question and the requirements of the current analysis.

There exist no general guidelines as to how game metrics tracking software should output the data, however, the few cases that are published (many systems being industry secrets) generally show log files of raw data transferred to a database system, from where they are extracted, analyzed and visualized:

00000: (x, y, z), (t), (w°), (λ°, φ°), (m), (i)
00001: (135, 451, 00), (0079), (111°), (100°, 75°), (i1)
00002: (155, 422, 05), (0148), (126°), (110°, 90°), (i3)
00003: (193, 420, 10), (0210), (141°), (180°, 110°), (w3)

This highly simplified version of the output provided by the IO Interactive metrics tracking software shows the recording of specific set of triggered metrics, providing a unique code for each, and recording the location of the PC, time since recording started; the camera angle and the type of metric recorded (last parenthesis). At ITU and IOI, *ArcView GIS* (Geographical Information System) is utilized to analyze all metrics with a spatial component, e.g. movement. In a GIS, data sets are added as layers on top of a map, and calculations can be performed either along layers or across, providing a very powerful tool for the analysis of spatial data and locating patterns of play.

Transforming the rough metrics data output from a metrics tracking software package to useful reports by Quality Assurance, researchers and designers, requires a series of potentially iterative steps [8]. The process involves obtaining the raw data, parsing them to provide structure and filtering, analyzing them to locate patterns or results required. This is followed by representation of the data in a visual model, which is iteratively defined to improve visual quality (and ideally interactivity).

To show an example of a simple visualization of metrics, the positioning of the entity or entities controlled by the player within the virtual game world will be considered (Figure 1). Navigation metrics is one of most useful group of metrics to record in games featuring 3D virtual environments, because it in these games typically is possible to track player progression through the game via the location of their avatars (player representation in the virtual environment). Navigation metrics are typically recorded as frequency metrics, e.g. position recorded every three seconds, but can also be triggered. A diverse range can be defined depending on the context, including movement as a function of time, movement modifiers (different types of movement, which depend on the game in question, for example: "still", "crouch", "lie down", "run", "walk", "teleport" or "fly" in the MMORPG *World of Warcraft.*), speed and direction. As noted in the output example above, tracking navigation metrics can include recording camera view type, angle and direction relative to the vector of character movement. This provides information about what the player is looking at relative to the character.

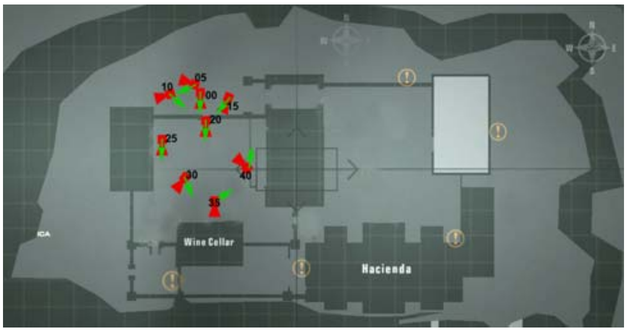

Figure 1: Tracking camera angle in *Hitman: Blood Money*. Green arrows indicate the vector of character movement; red icons indicate camera direction in the horizontal plane. (Figure created by A. Canossa, *IO Interactive*).

5. THE ROAD AHEAD

User experience is challenging to measure correctly and in detail within productivity software or hardware (e.g. mobile phones, machinery) [15], however, the experience arising from playing digital games is possibly even more difficult to get a handle on, because of the complex patterns of interaction between players and games. Unlike productivity applications, games do not exist to assist people solving tasks, but rather to entertain while the user is solving tasks that are not engineered to be as easy to solve as possible.

IO Interactive (IOI), the largest game development company in Denmark, has spent more than a year developing a comprehensive metrics tracking tool that is able to record virtually any user-initiated event in their as well as other games. One of the goals of the current collaboration between the *IT University Center for Computer Games Research* (ITU) and *IO Interactive* will be utilizing this tool to investigate how metrics can inform game design in practice, as well as in addressing key research-oriented issues such as whether players conform to specific patterns of play across different games or not.

What game experience is, and the components it consists of remain debated, however, some aspects such as emotional feedback and level of arousal can be measured using biometrics or qualitative methods. Hitherto, a major problem has been obtaining high-resolution, objective and quantitative data about player-game interaction, with which to correlate these measures of experience. Game metrics offer a tantalizing solution to this core problem.

The principal approach of the ITU-IOI project will be to combine metrics and biofeedback tracking, while conducting test sessions on a range of digital games at IO Interactive and the ITU. As discussed above, game metrics can inform what players are doing in the virtual environments and biofeedback how this affects them biophysically (and if e.g. measures of EEG were included, neurologically). However, traditional qualitative methods such as surveys and interviews are necessary in order to move beyond the "*what*" and the "*how*" and figure out the "*why*": Why do players find specific game design elements or events in the game storylines particularly arousing, frightening or fun? There exist a range of surveys for mapping some of the core components of game experience [e.g. 12], which will be integrated both during and following game sessions, following principles of [13]. Preliminary empirical work will be run in the fall 2008, and early results are expected to be ready for the NordCHI workshop in October.

6. ACKNOWLEDGEMENTS
The author would like to extend his warmest gratitude to colleagues at IO Interactive, notably Dr. Alessandro Canossa and User Research Manager Janus Rau Sørensen.

7. REFERENCES

[1] Blythe, M. A., Overbeeke, K., Monk, A. F. Wright, P.C. 2004. *Funology – from usability to enjoyment.* Springer Publishers.

[2] Bryant, J. and Vorderer, P. 2006. *Psychology of Entertainment*. Lawrence Erlbaum Associates.

[3] Cowley, B., Charles, D., Black, M. and Hickey, R. 2008.Toward an Understanding of Flow in Video Games. *ACM CiE*, 6(2), article 20.

[4] Csikszentmihalyi, M. 1991. *Flow: The Psychology of Optimal Experience*. Harper Perennial.

[5] Dibbel, J. 2006. *Play Money: Or, How I Quit My day Job and Made Millions Trading Virtual Loot*. Basic.

[6] DeRosa, P. 2007. Tracking Player Feedback to Improve Game Design. *Gamasutra*, August 7, 2007.

[7] Dornhege, G. 2007. *Toward Brain-Computer Interfacing*. MIT Press.

[8] Fry, B. 2008. Visualising Data. O´Reilly Media, Inc.

[9] Gilleade, K. M. and Dix, A. Using frustration in the design of adaptive videogames. In *Proc. of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology***,** 228-232.

[10] Gilleade, K., Dix, A. and Allanson, J. Affective Videogames and Modes of Affective Gaming: Assist Me, Challenge Me, Emote Me. In *Proceedings of the 2005 DiGRA Conference*.

[11] Ijsselsteijn, W., de Kort, Y., Poels, K., Jugelionis, A. and Bellotti, F. Characterizing and Measuring User Experiences in Digital Games. In *Proceedings of the ACE Conference 2007* (ACM Publishers).

[12] Isbister, K. *Better Game Characters by Design: A Psychological Approach.* Morgan Kaufman, 2006.

[13] Kim, J. H., Gunn, D. V., Schuh, E., Phillips, B. C., Pagulayan, R. J. and Wixon, D. 2008. Tracking Real-Time User Experience (TRUE): A comprehensive instrumentation solution for complex systems. In *Proceedings of CHI 2008*, 443-451.

[14] Koster, R. 2005. *A Theory of Fun*. Paraglyph Publishers.

[15] Kuniavsky, M. 2003. *Observing the User Experience: A Practitioner's Guide to User Research*. Morgan Kaufmann Publishers.

[16] Mandryk, R. L., Atkins, M. S. and Inkpen, K. M. A Continuous and Objective Evaluation of Emotional Experience with Interactive Play Environments. *Proc. of CHI 2006: Novel methods: Emotions, Gestures, Events,* ACM Press.

[17] Pagulayan, R. J., Keeker, K., Wixon, D., Romero, R. L. and Fuller, T. 2003. User-centered design in games. In *The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications*, 883-906. Lawrence Erlbaum Associates.

[18] Salen, K. and Zimmerman, E. 2003. *Rules of Play – Game Design Fundamentals*. MIT Press.

[19] Schultz, C. P. 2005. *Game Testing All in One*. Course Technology PTR.

[20] Sweetser, P. and Wyeth, P. GameFlow: A Model for Evaluating Player Enjoyment in Games. *ACM Computers in Entertainment*, 3(3), Article 3A (2005).

[21] Thompson, C. Halo 3: How Microsoft Labs Invented a New Science of Play. *Wired Magazine*, 15(9) (2007).

[22] Yannakakis, N., Hallam, J. Capturing Player Enjoyment in Computer Games. In *Advanced Intelligent Paradigms in Computer Games 2007*, 175-201.

[23] Yannakakis, N., Hallam, J. 2006. Towards optimizing entertainment in computer games, *Applied Artificial Intelligence*, 21, 933-971.