Methods for Evaluating Gameplay Experience in a Serious Gaming Context

Lennart Nacke
University of Saskatchewan
Lennart.Nacke@acm.org

Anders Drachen
Dragon Consulting
andersdrachen@gmail.com

Stefan Göbel
TU Darmstadt
stefan.goebel@kom.tu-darmstadt.de

Abstract

Gameplay experience (GX) is created during the process of player-game interaction, where this interaction has the goal to provide a motivating, fun experience for the player. Since GX is an important factor for the success of failure of a game, a formal classification of how to design for and evaluate GX is necessary. Using appropriate mechanisms for evaluation and measurement of GX allows the validation of good gameplay experiences. This paper presents an approach to formalize such evaluative methods and a roadmap for applying these mechanisms in the context of serious games. We first discuss related work of user experience (UX) and player experience models, based on which we propose a three-layer framework of GX. For each layer, a number of measurement methodologies are listed and our focus is put on physiological and technical metrics for game evaluation. Finally, we point out the potential use of this framework within the field of game-based learning and serious gaming for sports and health.

GAMEPLAY, AFFECT, PSYCHOPHYSIOLOGY, USER EXPERIENCE (UX), SERIOUS GAMING

Introduction

Over the past decade, an increasing amount of research interest has been directed toward the emotional and affective aspects of user experience (UX) that digital games provide. The evaluation of digital games has been a largely informal process in the past. However, the gaming industry is starting to adopt more formal techniques from human-computer interaction (HCI), especially UX, to evaluate their products. The caveat is that for game testing classic usability testing does not suffice, since its standard metrics (e.g., effectiveness in task completion or efficiency in error rate) are not directly applicable to all aspects of digital games (Pagulayan, Keeker, Wixon, Romero, & Fuller, 2003). Traditional usability metrics are still relevant, but they need to be adapted to digital games and supplemented with physiological and metrical assessment of gameplay experience (Mandryk, Atkins, & Inkpen, 2006; Nacke, 2009a; Nacke, Lindley, & Stellmach, 2008; Tychsen & Canossa, 2008). In the area of serious gaming (Gee, 2003; Michael & Chen, 2005; Prensky, 2000), a further extension is necessary to include *serious* aspects (advanced goals, such as training in a certain knowledge domain, apart from pure entertainment). For example, in the field of game-based learning, approaches such as evaluation schemes of computer-based learning arrangements have been developed (Bruder, Offenbartl, Oswald, & Sauer, 2004; Bruder, Offenbartl, Oswald, Sauer, & Sonnenberger, 2005) although they lack the inclusion of UX factors (and appropriate measurement methodologies) described in this paper.

Since recently much effort has been put into broadening usability concepts to investigate UX in terms of underlying principles and action plans for improving design, similar actions are now being taken to formalize playtesting methodology during game development (Nacke, 2009b; Pagulayan, et al., 2003; Sánchez, Zea, & Gutiérrez, 2009). However, there is still a lack of concrete advice for game industry and research practice about the taxonomical relationship of different methodologies for gameplay experience assessment.

We aim here at incorporating different measurements in a broad and simplistic framework, which should include methods already in use for game evaluation. We also want to focus on presenting emerging methodologies that combine physiological and technological metrics in the different layers of our framework that provide a technology-based, but user-centered approach to evaluating games during the development process. It is our hope that by establishing emerging game evaluation methodologies and framing them, this research will be useful for game developers and researchers alike.

Models of User Experience

UX has formed one of the cornerstones of HCI research in order of a decade (Law, Vermeeren, Hassenzahl, & Blythe, 2007). Thus, a variety of (A) *models* of UX and (B) considerations of *design principles* for UX, have emerged in this period of time with different foci, such as emotion, affect, experience, pleasure, hedonic qualities, and the like (Law, Roto, Hassenzahl, Vermeeren, & Kort, 2009). In this section, we show how this previous work forms the foundation for the framework that is presented here.

Aimed toward design practices, the fundamental work of Malone (1980) described what makes things fun to learn and presented a model of intrinsically motivating instruction. This has been influential in game design and game-based learning. It contains three categories of design guidelines: challenge, fantasy, and curiosity. The model is useful for the generation of design principles, but does not provide methods for evaluating UX.

Hassenzahl (2004) introduced a model (supplementing simpler and older models, e.g. Logan (1994)), which views UX from a designer and a user perspective making a distinction between the intended and apparent character of a product. Thus, he emphasizes the fact that there is no guarantee for designers to ensure their products are used or perceived as intended. The emotional personal response to a product is based on the situational context. Experience in his model is formed from the iconic value and prior memories the product triggers. Following his argumentation, a product can have pragmatic (e.g., utilitarian value) and hedonic (e.g., knowledge/skill stimulation, communication of identity, memory evocation) attributes. This model extends to games as well, since they also provide challenges, stimulation, and novelty to create personal value. However, games are primarily played for their hedonic value so that pragmatic use of play – which could be relevant for serious gaming – is often hidden underneath the pleasurable experience.

Garrett (2003) proposed a design model of UX for the web, with UX elements clustered on different layers of abstraction during the web development process. The core of the model consists of integrating UX considerations as the product development process moves from an abstract synergy to a concrete aesthetic surface. This notion of abstraction in modeling experience can be adapted to a game development context and inspired the abstraction dimension in the GX model that we will present in the next section.

Based on Maslow's motivational model of human needs (Maslow, 1943), Jordan (1999) proposed a hierarchical pleasure model of needs oriented toward design practices. In this model, pleasure follows from usability (which depends on functionality). He distinguishes four types of pleasure: physio-pleasure (e.g., evoked tactile and olfactory stimuli), socio-pleasure (e.g., evoked by relationships, society, personal status, or indicative of social identity), psycho-pleasure (e.g., cognitive and emotional reactions), and ideo-pleasure (e.g., aesthetics and ideological value).

Few researchers have tried to explicitly model GX or in a similar vein, *playability*. The latter is a somewhat nebulous concept, which is inconsistently used in the literature (Fabricatore, Nussbaum, & Rosas, 2002; Järvinen, Heliö, & Mäyrä, 2002; Sánchez, et al., 2009) although most sources utilize the term to refer to usability testing that is not focused on interface design, but for example game pacing, difficulty balancing, game mechanics and game story.

Arhippainen and Tähti (2003) described UX as being formed via user-product interaction only. However, they included considerations of the influence of social and cultural factors on UX. The interaction of these factors was viewed as contributing to UX. A primary drawback of the model is that it lacks reflection about UX in terms of systems design. This is a subject included by Kankainen (2003), who discusses user-centered methodologies for concept-level product design. UX is defined as a result of motivated action in specific contexts. Motivation, action, and content all form the vertices of a triangle that subsumes present experience. Thereby, he adds temporality to the UX debate, indicating that previous experiences shape present experiences, which will affect future experiences. The model focuses on informing content creation by evaluating users' motivational needs and action needs. Motivational needs are hard to assess with any kind of methodology, while action needs can be evaluated during gameplay if behavior and emotional responses of players are recorded.

Fernandez (2008) also worked with the temporal dimension of UX, proposing a model of GX which is built around temporal influences *before* (i.e., antecedents), *during* (i.e., processing) and *after* (i.e., consequences) the GX occurring during player-game interaction. The model considers "fun" to be the chief component of GX, which is constructed from emotional and cognitive player responses. Game evaluation should concentrate on evaluating these. The model is too specific to account for all instances of GX and playtesting in games, thus it is a good starting point for a broader, more inclusive GX model that has a development-centric perspective on evaluation and classification of GX.

A mapping of usability to playability to evaluate UX in entertainment systems by Sánchez et al. (2009) deconstructed playability, integrating methodological considerations from the game development process. They propose six "facets of playability": Intrinsic, mechanical, interactive, artistic, intrapersonal/personal, and interpersonal/social playability. However, this model has two major limitations: 1) It is not described what process that was used to derive the game playability concepts from desktop usability definitions; 2) The model introduces unclear concepts of playability that are in turn defined using other conceptual descriptions, which are empirically unsupported.

Qualitative playability evaluation in other research studies provides criteria with which a product's gameplay or interaction can be evaluated. Järvinen et al. (2002) proposed four components of playability each relating to formal aspects (i.e., functional) and informal aspects (i.e., experiential) of games. The challenge for qualitative playability evaluation is that most of them rely on expert evaluations.

One of the few practical approaches to playability evaluation is the use of sets of design heuristics. A few sets have been proposed for digital games. For example, Desurvire, Caplan, and Toth (2004) discussed four game heuristic categories: game play, game story, game mechanics, game usability. Korhonen and Koivisto (2006) presented playability heuristics modules for game usability, mobility and gameplay. Heuristics-based evaluation permits the development of a detailed understanding of the conceptual and individual factors determining gameplay, with the limitation that the approach relies on experience of the evaluator.

Fabricatore et al. (2002) used grounded theory to create a qualitative model of GX, which focuses on utilizing the opinions of players to develop design guidelines, which conceptualize individual playing preferences. A substantial body of work within the game studies field has been focused on conceptualizing and analyzing theoretical constructs, which are likely part the gameplay experience, such as flow (Cowley, Charles, Black, & Hickey, 2008; Sweetser & Wyeth, 2005), immersion (Ermi & Mäyrä, 2005; Jennett, et al., 2008), frustration (Gilleade & Dix, 2004), enjoyment (Gajadhar, de Kort, & IJsselsteijn, 2008; Klimmt, 2003), presence (Slater, 2002) and spatial presence (Wirth, et al., 2007), and so forth. The focus of this article is on providing a broad categorical model of GX, so that a detailed discussion of all these experiential constructs is beyond its scope.

Nacke (2009b) proposed a practice-oriented model that focused on describing playtesting in game development, classifying game usability into evaluations of technology (i.e., system quality), player (i.e., gameplay quality), and community (i.e., social quality). Analysis of the technology is handled by quality assurance team, player analysis by a UX team, and community analysis is done in sociological studies. The model comes close to Garret's (2008) description of UX design for the web, and is organized in layers similarly to the approaches in UX that are described above. It provides a starting point for the abstraction dimension of the model proposed in the next section.

Three important threads of UX research are apparent in the currently available literature: 1) The requirement for addressing human needs beyond the instrumental; 2) Affective and emotional aspects of interaction, and 3) The nature of UX itself (Hassenzahl & Tractinsky, 2006).

This provides two motivations for the framework presented here: (1) Providing a frame for technological methods that enable behavioral modeling of player-game interaction and (2) Providing a frame for evaluation of affect and cognition of player-game interaction. Different UX models and approaches for projecting UX in games have been discussed. However, none of these has taken an *inclusive* theoretical stance with focus on the game system, the player, and the context of play – especially in serious gaming environments such as game-based learning or games for sports and health, where the context is key to the training experience. Similar to product experience in general, GX is very complex, since player personas and profiling (also including individual background knowledge for educational games or the health and fitness status of a player being relevant for sports and health games); individual expectations based on past experiences are all factors contributing to it. Existing models also do not include recommendations about which methods are useful for measuring within different frames of UX or GX. Our model provides such recommendations within classifying frames of GX.

Measuring Gameplay Experiences

The digital game development process is usually iterative and product-focused. Thus, testing of game systems has classically been carried out by quality assurance groups with a focus on finding bugs in the software and has been synonymous with assuring technical quality of the digital game. Playtesting with user-focus on the player has long been performed with a high degree of informalism (e.g. recruiting testers from within the game development studio). Today we see playtesting adopt strategies from HCI and usability for developing inclusive player-game evaluation instrumentation (Isbister & Schaffer, 2008; Kim, et al., 2008; Pagulayan, et al., 2003).

There are three methodological categories for experiences that surround digital games: the quality of the product (game system experience), the quality of human-product interaction (individual player experience), and the quality of this interaction in a given social, temporal, spatial or other context. All of these qualities will determine different layers of gameplay experience over time (since time is an important influence factor, which is subsumed in all the frames) and all of them can be assessed during the game development process. Figure 1 shows how these methodological gameplay experience layers interact in game development. For example, what the developer implements in the game system affects player experience. The contextual effects on player experience are diverse, such as presence of other players, physical health conditions, time of day, and many others.

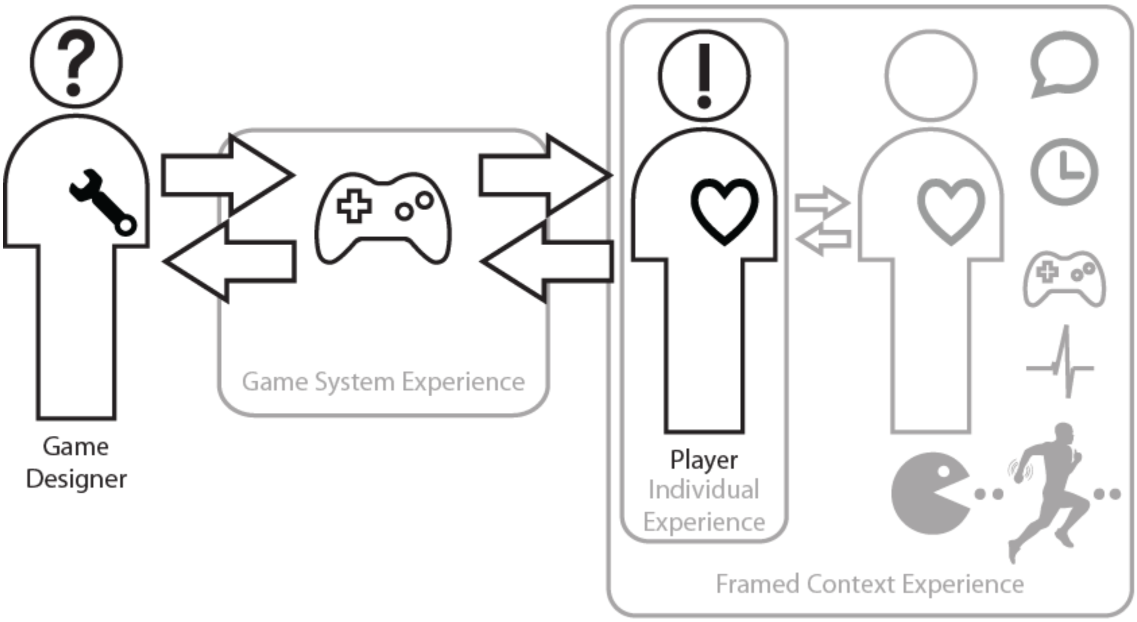

Figure 1. Three methodological frames of gameplay experience in the game development process. For example, game system experience methods are concerned with functional testing of the game; player experience methods ideally use sensor technology (or usability and playtesting) to assess emotion and enable player-game interaction, and finally logging metrics methods (among others) enable assessing game context experience.

Game developers currently have primarily influence on game system experience by refining and testing the game software and by balancing the game system variables. Systematic balancing of gameplay requires either a very good knowledge of the preferences of a targeted user base or the application of scientific methods to evaluate user preferences, emotions, and behavior. This affects the layer of individual player experience, which models the reception and effects, the game mechanics, dynamics, and aesthetics have on a player. Thus, game developers can only indirectly influence this layer of game experience, for example by player modeling in the game system (Drachen, Canossa, & Yannakakis, 2009) or by providing adaptive technologies (e.g., macro and micro adaptation) in digital educational games (Göbel, Mehm, Radke, & Steinmetz, 2009; Peirce, Conlan, & Wade, 2008). This is also the layer where affective studies allow a prominent assessment of player emotion and cognition using psychophysiological methodology. Finally, the individual experience is also affected by the context in which playing happens. This can for example relate to playing in a social context that might amplify emotions (e.g., playing a party/exergame such as *Wii Sports* in a competitive multi-player scenario vs. playing a health game such as *Wii Fit* in a stand-alone mode), it might also relate to the temporal change of experience if a game is played more than one time and affection becomes intertwined with the experiential memory of playing a certain part or sequence in the game and the actual experience, which is elicited when the game is played again. The experience is then framed in a context. This context can only be marginally influenced by game developers through providing additional tie-in experiences[1]. In the following sections, we will look at methodologies that operate on the three different levels of the framework.

Assessing Game System Experience

Methodologies for assessing game system experience come from regular software and traditional game testing. They are often explicitly included in the game development process (at least by larger companies) and make sure the functional level on which the game system operates is correct. We will only briefly review these. Unit testing (i.e., automated testing of the program code), stress testing (i.e., testing of software or hardware limitations), soak testing, compatibility testing, regression testing, bug tracking, localization testing, open beta-testing, gameplay metrics (i.e., event- and location-based metrics allow the tracking of user behavior in the game, which can then feedback into the design process to balance out a certain level based on statistical data).

Assessing Individual Player Experience

An important issue preventing successful evaluation of digital games is the inability to successfully cater to individual emotions. New sensor technology and behavioral tracking enable us to model and assess player cognition and emotion during gameplay. Some examples of novel methodologies used here include:

- **Psychophysiological player testing**. These are controlled measures of gameplay experience usually deployed in a laboratory with the benefit of covertly assessing physical reactions of players.
    - Electromyography (**EMG**) is a measurement technology for recording the electrical activation of muscles. Basic emotions are well reflected in facial expressions. This allows a mapping of emotions in the valence dimensions of the circumplex model of affect (Russell, 1980).
    - Measurement of electrodermal activity (**EDA**) is one of the easiest and therefore most commonly used psychophysiological methods. The measured increased sweat gland activity is directly related to physical arousal.

[1] For example, the Xbox Live Pub Games of Fable II are a tie-in to the full game. They provide people that like the casual experiences of these games with a different experience than playing the full game, potentially evoking curiosity for the full game.

   - Electroencephalography (**EEG**) requires the participant to wear scalp electrodes. Brain waves are usually described in terms of frequency bands, such as alpha (e.g., 8-14 Hz).
   - Functional magnetic resonance imaging (**fMRI**), positron emission tomography (**PET**) and functional near-infrared spectroscopy (**fNIR**) are other non-invasive techniques for measuring brain activation. The former two have major limitations in deployment and are hardly used in UX research, while fNIR has recently received more attention from the HCI community.

- **Eye Tracking**. Eye Trackers measure the saccades (fast movements) and fixations (dwell times) of human gaze. Due to the relationship between eye fixations and attentional focus, we are able to infer and visualize cognitive and attentional processes in virtual environment exploration (Stellmach, Nacke, & Dachselt, 2010)
- **Persona Modeling**. "Play-personas" are partly data-driven and constructed. Persona models can be compared with user behavior metrics and prompt changes in the game design (Drachen, et al., 2009; Tychsen & Canossa, 2008).
- **Game Metrics Behavior Assessment**. Instrumentation data ideally log any action the player takes while playing, such as input commands, location, events, or interaction with in-game entities. As an analysis tool, metrics supplement existing methods of game user research by offering insights into how people are actually playing the games under examination (Tychsen & Canossa, 2008).
- **Player Modeling**. Research based on AI is using neural networks and cognitive theories to model and react to player behavior, with the goal to develop adaptive games. Models of players based on behavior and responses to different in-game situations, form the basis for how the game should adapt in real time (Drachen, et al., 2009).
- **Qualitative interviews and questionnaires**. Semi-quantitative and qualitative approaches traditionally form the basis for user-feedback gathering in game development. Surveys have been focused on the enjoyment-aspects of UX in games, but recent developments have included dimensions such as tension, frustration, or negative affect. Using surveys during natural breaks in the gameplay action is usually the preferred method of deployment.
- **RITE Testing**. Rapid Iterative Testing and Evaluation was developed by Microsoft Game User Research (Medlock, Wixon, Terrano, Romero, & Fulton, 2002). The approach specifies data analysis after each participant or at the end of the testing day with changes to the interface or the game design being made rapidly after a solution is found. This allows iterative improvement of game designs during the development process.

Assessing Player Context Experience

Playing context is not often evaluated by game development studios, but has been the focus of research efforts, mainly from empirical and sociological perspectives. This is especially common in testing mobile games, where the context of the game has a large influence on how the game is perceived (Korhonen & Koivisto, 2006). Examples of methodologies used here include:

- **Ethnography**. Ethnographic methods attempt to record practices of a certain population acknowledging the impossibility for the researcher to be a transparent observer but instead treasure the impact that the act of observing has on the studied population.
- **Cultural debugging**: Testing conducted to assess how and if culturally arbitrary conventions are understood in different contexts. For example in Deus Ex 3, a receptionist was not perceived as such outside of the US because of cultural conventions.
- **Playability Heuristics**. Playability heuristics can be implemented quickly and cheaply into the game development process. There a few sets of specialized heuristics for use in game development (Desurvire, et al., 2004; Korhonen & Koivisto, 2006). Expert reviews with heuristics have been presented for action games, based on technical game review scores, and for game-based learning applications. The main benefits of heuristic methods are that they are time and cost efficient.
- **Qualitative interviews and questionnaires**. Used to assess context and social impact on individual player experience in a similar capacity as in the UX assessment at the individual player experience level.
- **Multiplayer game metrics.** Similar as for player experience, the social experience can be modeled using gameplay metrics to study the interaction of several players depicted the interaction log.

Applying the Framework to Serious Games

Serious games have been used before as motivators for learning (Gee, 2003; Prensky, 2000) and for healthy, sportive activity (Baranowski, Buday, Thompson, & Baranowski, 2008; Wiemeyer, 2009). The sensor technology and the psychophysiological (measurement) methods discussed within our proposed framework provide great possibilities in such application scenarios, both (1) to enhance adaptation and personalization of existing games and systems during play and (2) to evaluate the effectiveness of serious games in general. The following two examples come from ongoing research activities of the serious gaming group at TU Darmstadt and we will apply our framework to these examples.

- In *80Days*[2] different game modes (relaxed, driven and hectic version) are used to match player characteristics (taken from a player questionnaire administered before the educational game starts) as good as possible. For example, players familiar with action games receive the arousing or hectic version with exciting music, (artificial) time pressure and alien opponents chasing them by default, whilst casual gamers without an affinity for (action) gaming and interactive media play a relaxed version at first. Then, during gameplay, individual player experience is evaluated by classical event logging mechanisms and user interactions with the system. By using the GX framework from this paper, we can decide to add some of the above methods for evaluating individual player experience. For example,

[2] For more information about the game see http://www.eightydays.eu

we will check the emotional state of players and the appropriateness of particular game modes or of switching modes during play. Similar, the assessed player experience might be considered in real time during gameplay to determine how an educational game continues during play (i.e., macro adaptation; sequencing of learning objects and learning units/game levels) or how a learning unit/object is presented (i.e., level of interactivity, use of digital media; interaction templates, more specifically micro adaptation).

- In *motivation60+*[3] the overall aim is to provide and to combine sensor technology and game-based concepts in order not only to monitor the health status of (elderly) users, but also to motivate the *silver generation* to do sports regularly, in a sustainable manner. Here, in addition to sensor technology for the measurement of vital parameters such as the heart rate and other psychophysiological methods would provide a motivating insight into the affective player experience and "soft factors" such as emotion, fun, arousal, stress, and others. These GX parameters are highly relevant for the success or failure of game design and subsequently for a long-term, sustainable motivation and effectiveness of sports/health games.

The individual player experience is important in serious gaming, since this frame is where an effect of the serious game on behavioral change in the player can be witnessed. It depends on the kind of serious game that should be tested, which evaluative UX measure is preferred, but our recommendation is that affective measures should be used when probing for effectiveness of game-based learning applications for example.

Conclusion and Future Work

In brief, UX evaluation of gameplay can be categorized in three GX frames. Game system experience methods work on the functional level of assessing the functional capacity of the game system, game engine, or level data to balance the game design. Player experience methods evaluate the emotional or cognitive impact that a game or certain events or entities in a game have on a player, ideally by using quantifiable, objective, and physiological methods. Finally, context experience methods are suited for studying the interaction of player in a co-located or co-present game environment, taking into account the sociological impact of game mechanics and player behavior. Ideally, those methods allow balancing and tuning of the game during and after game development to improve the experience and contextual or personal value that a game can provide to its players.

The individual player experience is especially important in serious gaming, since this frame is where we want to use a serious game to foster learning and behavioral changes. It depends on the kind of serious game that should be tested, which evaluative UX measure is chosen, but our recommendation is that affective measures are ideal for evaluating the effectiveness of motivation to play game-based learning applications or games for sports and health. While these measures might focus on emotional assessment, we assume that a link between positive emotion and long-term storage and recall of information in the brain exists, which future studies will investigate using the methods presented in this framework.

Further research will be investigating qualitative and quantitative evaluation studies for serious games domains. In a first step, comparative studies of *80Days* with different

[3] For more information about the project see http://www.motivotion.org

game modes and classical evaluation mechanisms (questionnaires, interviews, assessment) will be carried out to receive valid feedback about usage effects of different game modes and the overall effectiveness of learning. Finally, another goal is to find a correlation between fun, identification in games and (successful) learning. Similar, in *motivotion60+*, the aim is to correlate the intensity and duration of movements in exergaming with the attractiveness and fun of a sports or health game.

References

Arhippainen, L., & Tähti, M. (2003). *Empirical Evaluation of User Experience in Two Adaptive Mobile Application Prototypes* MUM 2003. Proc. of the 2nd Int. Conf. on Mobile and Ubiquitous Multimedia, Norrköping, Sweden.

Baranowski, T., Buday, R., Thompson, D. I., & Baranowski, J. (2008). Playing for real. Video games and stories for health-related behavior change. *American Journal of Preventive Medicine, 34*(1), 74–82.

Bruder, R., Offenbartl, S., Oswald, K., & Sauer, S. (2004). *Third party certification of computer-based learning environments.* EISTA: Int. Conf. on Education and Information Systems: Technologies and Applications, Orlando, Florida.

Bruder, R., Offenbartl, S., Oswald, K., Sauer, S., & Sonnenberger, J. (2005). Qualität von computergestützten Lernumgebungen beurteilen: ein ganzheitliches Konzept [Judging the quality of computer-based learning environments: a wholistic concept]. *Thema Forschung: E-Learning, 1*(2005), 62-65.

Cowley, B., Charles, D., Black, M., & Hickey, R. (2008). Toward an understanding of flow in video games. *Computers in Entertainment, 6*(2), 1-27.

Desurvire, H., Caplan, M., & Toth, J. A. (2004). *Using heuristics to evaluate the playability of games.* CHI '04 extended abstracts, Vienna, Austria.

Drachen, A., Canossa, A., & Yannakakis, G. N. (2009). *Player Modeling using Self-Organization in Tomb Raider: Underworld.* Proc. of IEEE CIG 2009, Milano, Italy.

Ermi, L., & Mäyrä, F. (2005). *Fundamental Components of the Gameplay Experience: Analysing Immersion.* Proc. of DiGRA.Vancouver, BC, Canada.

Fabricatore, C., Nussbaum, M., & Rosas, R. (2002). Playability in video games: a qualitative design model. *Human-Computer Interaction, 17*(4), 311-368.

Fernandez, A. (2008). Fun Experience with Digital Games: A Model Proposition. In O. Leino, H. Wirman & A. Fernandez (Eds.), *Extending Experiences: Structure, Analysis and Design of Computer Game Player Experience* (pp. 181-190). Rovaniemi, Finland: Lapland University Press.

Gajadhar, B. J., de Kort, Y. A. W., & IJsselsteijn, W. A. (2008). *Shared Fun Is Doubled Fun: Player Enjoyment as a Function of Social Setting.* Proc. of Fun and Games 2008, Eindhoven, The Netherlands, Oct. 20-21, 2008.

Garrett, J. J. (2003). *The Elements of User Experience: User-Centered Design for the Web*. Berkeley, CA, USA: New Riders.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York, NY, USA: Palgrave Macmillan.

Gilleade, K. M., & Dix, A. (2004). *Using frustration in the design of adaptive videogames.* Proc. of ACE 2004, Singapore.

Göbel, S., Mehm, F., Radke, S., & Steinmetz, R. (2009). *80days: Adaptive Digital Storytelling for Digital Educational Games.* Proc. of 2nd Int. Workshop on Story-Telling and Educational Games (STEG'09), Aachen, Germany.

Hassenzahl, M. (2004). The thing and I: understanding the relationship between user and product. In M.Blythe, C. Overbeeke, A. F. Monk & P. C. Wright (Eds.), *Funology: From Usability to Enjoyment* (pp. 31-42). Norwell, MA, USA: Kluwer Academic Publishers.

Hassenzahl, M., & Tractinsky, N. (2006). User experience - a research agenda. *Behaviour & Information Technology, 25*(2), 91-97.

Isbister, K., & Schaffer, N. (2008). *Game Usability: Advice from the experts for advancing the player experience*. Burlington, MA: Morgan Kaufmann.

Järvinen, A., Heliö, S., & Mäyrä, F. (2002). *Communication and Community in Digital Entertainment Services* (Prestudy Research Report). Hypermedia Laboratory, University of Tampere.

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., et al. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies, 66*, 641-661.

Jordan, P. W. (1999). Pleasure with Products: Human Factors for Body, Mind and Soul. In W. S. Green & P. W. Jordan (Eds.), *Human Factors in Product Design: Current Practise and Future Trends* (pp. 206-217). UK: Taylor & Francis.

Kankainen, A. (2003). *UCPCD: User-centered product concept design*. Proc. of Conf. on Designing for User Experiences, San Francisco, CA, USA.

Kim, J. H., Gunn, D. V., Schuh, E., Phillips, B., Pagulayan, R. J., & Wixon, D. (2008). *Tracking real-time user experience (TRUE): a comprehensive instrumentation solution for complex systems*. Proc. of CHI 2008, Florence, Italy.

Klimmt, C. (2003). *Dimensions and Determinants of the Enjoyment of Playing Digital Games: A Three-Level Model*. Proc. of Level Up: DiGRA, Utrecht.

Korhonen, H., & Koivisto, E. M. I. (2006). *Playability heuristics for mobile games*. Proc. of 8th conf. on HCI with Mobile Devices and Services, Espoo, Finland.

Law, E., Vermeeren, A. P. O. S., Hassenzahl, M., & Blythe, M. (2007). *Towards a UX manifesto*. Proc. of HCI 2008: People and Computers XXI: HCI...but not as we know it - Volume 2, Swinton, UK.

Law, E. L.-C., Roto, V., Hassenzahl, M., Vermeeren, A. P. O. S., & Kort, J. (2009). *Understanding, scoping and defining user experience: a survey approach*. Proc. of CHI 2009, Boston, MA, USA.

Logan, R. J. (1994). Behavioral and Emotional Usability: Thomson Consumer Electronics. In M. Wiklund (Ed.), *Usability in practice: how companies develop user-friendly products* (pp. 59-82). San Diego, CA, USA: Academic Press.

Malone, T. W. (1980). *What makes things fun to learn? heuristics for designing instructional computer games*. Proceedings of the 3rd ACM SIGSMALL Symposium, Palo Alto, CA, United States.

Mandryk, R. L., Atkins, M. S., & Inkpen, K. M. (2006). *A Continuous and Objective Evaluation of Emotional Experience with Interactive Play Environments*. Proc. of CHI 2006, Montréal, Québec, Canada.

Maslow, A. H. (1943). A Theory of Human Motivation. *Psychological Review, 50*(4), 370-396.

Medlock, M. C., Wixon, D., Terrano, M., Romero, R. L., & Fulton, B. (2002). *Using the RITE method to improve products: A definition and a case study*. Proc. of UPA 2002, Orlando, Florida, USA.

Michael, D. R., & Chen, S. L. (2005). *Serious games: Games that educate, train, and inform*. Boston, MA: Thomson Course Technology.

Nacke, L. (2009a). *Affective Ludology: Scientific Measurement of User Experience in Interactive Entertainment.* Ph.D. Thesis, Blekinge Institute of Technology, Karlskrona. Available online at: http://affectiveludology.acagamic.com

Nacke, L. (2009b). *From Playability to a Hierarchical Game Usability Model.* Proc. of Future Play @ GDC Canada, Vancouver, BC, Canada.

Nacke, L., Lindley, C., & Stellmach, S. (2008). *Log who's playing: psychophysiological game analysis made easy through event logging.* Proceedings of Fun and Games 2008, Eindhoven, The Netherlands.

Pagulayan, R., Keeker, K., Wixon, D., Romero, R. L., & Fuller, T. (2003). User-centered design in games. In J. A. Jacko & A. Sears (Eds.), *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications* (pp. 883-906). New York, NY, USA: L. Erlbaum Associates Inc.

Peirce, N., Conlan, O., & Wade, V. (2008). *Adaptive Educational Games: Providing Non-invasive Personalised Learning Experiences.* 2nd IEEE Int. Conf. on Digital Games and Intelligent Toys Based Education, Banff, Canada.

Prensky, M. (2000). *Digital Game-Based Learning.* New York, NY, USA: McGraw-Hill.

Russell, J. A. (1980). A Circumplex Model of Affect. *Journal of Personality and Social Psychology, 39*(6), 1161-1178.

Sánchez, J. L. G., Zea, N. P., & Gutiérrez, F. L. (2009). *From Usability to Playability: Introduction to Player-Centred Video Game Development Process.* Proc. of 1st Int. Conf. HCD 2009, San Diego, CA, USA.

Slater, M. (2002). Presence and the sixth sense. *Presence: Teleoperators and Virtual Environments, 11*(4), 435-439.

Stellmach, S., Nacke, L., & Dachselt, R. (2010). *Advanced Gaze Visualizations for Three-dimensional Virtual Environments.* Proc. of ETRA 2010, Austin, TX.

Sweetser, P., & Wyeth, P. (2005). GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE), 3*(3), 3-3.

Tychsen, A., & Canossa, A. (2008). *Defining personas in games using metrics.* Proceedings of the 2008 Conference on Future Play: Research, Play, Share, Toronto, Ontario, Canada.

Wiemeyer, J. (2009). *Serious games in prevention and rehabilitation.* International conference „Physical education and sport in research 2009 – Aging and physical activity", Rydzyna, Poland.

Wirth, W., Hartmann, T., Böcking, S., Vorderer, P., Klimmt, C., Schramm, H., et al. (2007). A Process Model of the Formation of Spatial Presence Experiences. *Media Psychology, 9*(3), 493-493.