HAL
open science

Heuristics for Video Games Evaluation: How Players Rate Their Relevance for Different Game Genres According to Their Experience

Florentin Rodio, J. M. Christian Bastien

To cite this version:

Florentin Rodio, J. M. Christian Bastien. Heuristics for Video Games Evaluation: How Players Rate Their Relevance for Different Game Genres According to Their Experience. 25ème conférence francophone sur l'Interaction Homme-Machine, IHM'13, AFIHM, Nov 2013, Bordeaux, France. 10.1145/2534903.2534915 . hal-00877287v2

HAL Id: hal-00877287

https://inria.hal.science/hal-00877287v2

Submitted on 6 Nov 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Heuristics for Video Games Evaluation: How Players Rate Their Relevance for Different Game Genres According to Their Experience

Florentin Rodio
University of Lorraine
Laboratoire PErSEUs (EA 7312)
57006 Metz (France)
florentin.rodio8@etu.univ-lorraine.fr
Alcatel-Lucent Bell Labs France, Route de Villejust, 91620 Nozay, France
florentin.rodio@alcatel-lucent.com

J. M. Christian Bastien
University of Lorraine
Laboratoire PErSEUs (EA 7312)
57006 Metz (France)
christian.bastien@univ-lorraine.fr

ABSTRACT
This paper reports a study evaluating the relative relevance of evaluation heuristics for different game genres and players. 120 players (amateur or e-Sport players) were invited to assess the relevance of 47 heuristics for different game genre (Real Time Strategy, Massively Multi-player Online Role-Playing or First-Person Perspective Shooting). Results show that the relevance of heuristics varies according to game genres and player types. These results have implication for both the design and evaluation processes of games. Implications of these results and future research directions are discussed.

Author Keywords
Video Games; Methodology; Evaluation; Game genres; Player typologies; Heuristics.

ACM Classification Keywords
H.5.2 [Information Interfaces and Presentation]: User Interfaces – graphical user interfaces, evaluation/methodology, theory and methods.

INTRODUCTION
Increased attention has been dedicated to the scientific study of video games given the economic appeal of the field. One of these concerns has been the improvement of evaluation methods, such as heuristics evaluation, a flexible, quick and suitable method for most development studios. This paper reports on a study aimed at assessing the relevance of heuristics as judged by different users' types and for different game genres.

By supplying evidences of heuristics relative relevancies, we believe that it will help designers focus on important aspects of the game, according to their game's genres or player's expectations, and to avoid wasting their time on issues that may not be important. In the next section, we present current inspection methods for video games, as well as the most common game and player taxonomies.

© ACM, 2013. This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in Actes de la 25ième conférence francophone sur l'Interaction Homme-Machine, 2013.
http://dx.doi.org/10.1145/2534903.2534915

Inspection methods for Video Games
There are two main approaches for the evaluation of games: the empirical and the analytical approach. Empirical evaluation methods are based on user's performances or opinions while the analytical evaluation methods are based on the examination of a product using a set of theories or models [16]. The latter have the advantage of being performed early in the development process and at a low cost [13]. In the field of video games, these analytical assessment methods are usually performed by the designers themselves or during beta testing. These evaluations however are lacking in rigor and tend to be accompanied by more structured inspection methods such as heuristic evaluation [13] or the use of ergonomic criteria [16]. Basically, a set of evaluators inspects the interface with respect to a small set of fairly broad principles, referred to as the "heuristics". In recent years, several sets of heuristics for video games design and evaluation have been proposed. They are based on dimensions such as usability [8,14], playability [6,8,10] or player experience [7,17].

Game genres
To categorize video games, we often refer to the notion of genre. A genre is a category of games with similar type of interaction (or "gameplay"). One of the most famous classifications is given by Callois [3], who classifies play into four categories: *Agon* (competition), *Alea* (chance), *Mimicry* (role playing), *ilinx* (vertigo). In this study, we used three subgenres that belong to three widely accepted genres in the video game industry: "strategy game", "action game" and "role-playing game" (Figure 1). Strategy games involve managing various resources and units to accomplish a given objective and require high decision-making skills. The goals of the strategy games could be either to build a city or to destroy an enemy base. It can be in real time (RTS) or

Figure 1. Examples of RTS, FPS and MMORPG game genres, respectively.

turn based. Action games are mostly based on players' dexterity and reflexes. It includes first-person perspective shooting games (FPS). Role playing games are focused on character development and its progression through a scenario. « Massively Multi-player Online Role-Playing Games » (MMORPG) are part of it, and enable the player to operate in a persistent world.

Typologies of players

Typologies of players are usually based on skill level, personality or motivation. To differentiate the players in terms of level, we generally refer to *casual gamers* and *hardcore gamers*. They are usually differentiated in terms of play time, « *meta-gaming* » involvement or motivation. Some hardcore gamers are also called pro-gamers or e-sportsmen. These players enjoy competition and play video games as high-performance sports. Other player models are based on personality or play motivation. Bateman and Boon [2] identifies four personality types: Killer, Achiever, Explorer and Socializer. Lazarro [11] outlines four main motivational factors of players: Hard Fun (personal triumph over adversity), Easy Fun (Curiosity), Altered State (Relaxation / excitement) and People Factor (social Interaction). Finally, Yee [18] identifies 10 online gaming motivations that fall into 3 higher-level categories related to achievement, social, and immersion motivations.

RESEARCH QUESTION

When approaching a field as rich as video games, it is essential to consider that heuristics' relevance can differ according to game genres and player types. Indeed, heuristics guide the design of game elements in a specific way, and, it is this particular balancing that establishes a *game play* proper to a specific game genre. From the perspective of existing inspection models, many authors are aware that game genres are an important factor guiding the choice and the weighting of evaluation criteria [10,12,15]. However, none of them have provided strong and formal evidences that heuristics relevancy differs according to game genres. Similarly, only Desurvire and Wiberg [5] have developed a list of principles specifically for a given player type, the casual player. But, similarly, no author has provided evidence that heuristics relevancy differs according to different player types. For this reason, the aim of this study was to assess whether heuristics have different level of relevance according to (1) game genres and (2) player types.

This study differs significantly from classical heuristics studies on three main points. First, this study does not focus on the validation of a particular set of heuristics. As Cockton et al. [4] have stated, we think that fundamental knowledge on inspection methods must be acquired before via controlled experimentation. It is especially true in video game inspection, since contextual variables such as game genres and player profiles seem to play a key role in game evaluation. Secondly, this study focuses solely on heuristics "relevance". Usually, to decide which issues should receive developer attention first, severity rating is used. However, other criteria have been considered in video game inspection to refine the prioritization of issues such as genre ratings [12]. Relevance indicator follows this trend but also benefit of a broader use: it can be seen as a generic tool to prioritize heuristics according to external criterion. Finally, this study is based on player judgment instead of usability expert. As other researcher [6,9], we think that taking into account real gamer experiences would improve the validity of game studies. Indeed, because video game heuristics are based more on playability and player experience, confronting target users to those heuristics would be very informative. The only likely bias could be to underestimate the relevance of some usability heuristics because players do not fully understand their scopes.

METHODOLOGY

Material

To test the aforementioned assumption, a representative set of heuristics was chosen, as well as a given taxonomy of game genre and player type. The taxonomy used was kept intentionally simple to optimize the statistical power of our tests and we did not intend to represent the whole diversity of existing genre or player types.

Category	Heuristics
Game Play	A1. The players find the game fun, with no repetitive or boring tasks
	B3. Easy to learn, harder to master
	C1. The game world reacts to the player and remembers their passage through it
	E1. The game supports a variety of game styles
Usability & Game Mechanics	B1. Game controls are consistent within the game and follow standard conventions
	C2. Provide appropriate audio/visual/visceral feedback (music, sound effects, controller vibration)
	E1. The game does not put an unnecessary burden on the player
	H3. Upon turning on the game, the player has enough information to begin play
	G1. Navigation is consistent, logical and minimalist
Miscellaneous	A1. There is an emotional connection between the player and the game world as well as with their "avatar"

Table 2. Examples of heuristics from the PLAY2009 model [6]

Three game genres were chosen: the *RTS*, *MMORPG* and *FPS*. These three genres have accounted for over 60% of the sales in PC market in 2010[1]. In addition, each genre possesses a high-level competitive scene. A representative game was then selected for each genre: Red Alert 3 for RTS, Call of Duty 4 for FPS and World of Warcraft for MMORPG. These games were chosen because they possess a large gaming community and are diverse enough to reach all players types.

Amateurs and *e-Sports* players were then chosen as participants. E-Sports players fulfill themselves in competition and high performance situations whereas amateur players tend to play more for relaxation and social interaction.

Among all inspection models available, the PLAY 2009 [4] has been chosen. Composed of 48 heuristics (only 47 are available in the literature[2]), this model is divided into three sections: *Game Play*, *Usability & Game Mechanics* and *Coolness / Entertainment / Humor / Emotional / Immersion* (renamed in *Miscellaneous*). Examples of heuristics can be seen in Table 2.This model covers both the usability and playability of games and was developed to cover three genres (Action Adventure, FPS, and RTS).

Participants and procedure

A participation call has been posted on 70 general, specialized and guild video games websites. Each player, that belongs to one of the following two categories (amateur or e-Sport player) and one of the three game genres communities (STR, MMORPG or FPS) was asked to fill in a survey and rate the relevance of 47 heuristics related to their dedicated game genre on a five point Likert-type scale. In addition, questions related to game practices were also asked. Categorization between amateur and e-sports players has been accomplished by questions informing on their competition practice. On 232 invitations, 154 questionnaires were completely filled in and 120 were selected for this study. 20 amateurs and 20 e-sport players were selected for each of the three game genres.

RESULTS

Sample characteristics

The sample of players is predominantly composed of male (98.3%). Fifty percent of the participants are students, 18.3% are employees, 12.5% senior executive, 11.6% without employment, and others are 7.5%. The mean age of the players is 20 years old and their game experience is about twelve years. E-Sport players have significantly more experience in competition than amateurs on the specified game (1.53 vs 0.02 years, $t(118) = 8.31$, $p < .001$), on games of same genre (3.38 vs 0.05 years, $t(118) = 12.14$, $p < .001$) and spend significantly more time per week in competition on the game (12.83 vs 0.03 Hours, $t(118) = 8.946$, $p < .001$). This result supports and justifies the initial categorization between those two types of players.

Heuristics relevancy

Players' responses to the 47 criteria were analyzed with a 3 (RTS vs FPS vs MMORPG) x 2 (Amateur vs e-Sport players) ANOVA. Sheffe's Post Hoc test was carried out to look for significant pairwise differences between conditions. In addition, binomial tests have been performed to test whether the proportion of significant differences on the 47 criteria differ significantly from chance.

Seven out of 47 heuristics' relevance differ statistically according to game genres ($p < .05$). To test whether this observed proportion (15%) was statistically different from the expected proportion of 5% (the conventional error used when performing inferential test), a binomial test $B(47, 0.05)$ was performed. This test is significant ($p < .01$). Therefore, the difference of heuristics' relevance found between game genres is significant. However, we cannot conclude that all observed differences are only

[1]Annual survey of the *ESA* http://www.theesa.com/facts/pdfs/ESA_EF_2011.pdf

[2] The 47 heuristics used in this study are available here: http://www.behavioristics.com/downloads/DesigningBetterGames-09HCI-Desurvire.pdf

due to the game's type as the average difference due to chance in this situation is 2.35 (σ = 1.49)[3]. Because of this result, we have chosen to illustrate trends only with the most conservative statistical difference for each type of game. Thus, we note a significant difference between game genres for the criterion « *GamePlay E2: The game is balanced with multiple ways to win* » ($F(2,117) = 7.67$, $p < .001$). A Sheffe test indicates that this difference is significant between RTS and FPS ($p < .01$) and between RTS and MMORPG ($p < .01$). Hence, game balancing and variety of possible strategies seem to be an important criterion for RTS. Similarly, there is a significant difference between game genres for the criterion « *GamePlay D3: The game gives rewards that immerse the player more deeply in the game by increasing their capabilities, capacity or for example, expanding their ability to customize* » ($F(2,117) = 5.07$, $p < .01$). A Sheffe test indicates that this difference is significant between MMORPG and FPS ($p < .05$). So, rewarding in the game seems to be a more important criterion in MMORPGs than in FPS. As a last example, there is a significant difference between game genre for the criterion « *Miscellaneous C1: The game uses humor well* » ($F(2,117) = 5.53$, $p < .001$). A Sheffe test indicates that this difference is significant between FPS and RTS ($p < .01$) and between FPS and MMORPG ($p < .05$). Hence, humor in FPS seems to be of minor importance.

14 out of 47 heuristics differ significantly according to player types ($p < .05$). To see if this observed proportion (30%) was statistically different from the expected proportion of 5%, a binomial test $B(47, 0.05)$ was also computed. This test is significant ($p < .001$). Therefore, the difference of heuristics' relevance found between player types is significant. Like the previous section, we have illustrated trends only with the most conservative statistical difference between the two types of player. First, all these significant differences show a lower relevance of the heuristics by e-Sports players compared to amateur players. Many of them are related to artificial intelligence (« *GamePlay B5: AI is balanced with the players' play* », $t(118) = 3.31$, $p < .001$;« *GamePlay B6: The AI is tough enough that the players have to try different tactics against it.* », $t(118) = 2.28$, $p < .01$; « *GamePlay E4: The game had different AI settings so that it was challenging to all levels of players, whether novice or expert players* », $t(118) = 4.09$, $p < .001$). Furthermore, we find that many of these criteria have high importance in single mode (« *Usability & Game Mechanics H5: All levels of players are able to play and get involved quickly and easily with tutorials, and/or progressive or adjustable difficulty level* », $t(118) = 3.04$, $p < .01$; « *Usability & Game Mechanics I1: Game story encourages immersion*», $t(118) = 2.92$, $p < .01$; « *GamePlay A3: The players should not lose any hard won possessions* », $t(118) = 2.46$, $p < .01$). Hence, all these heuristics seem to be of minor importance for e-Sports players.

[3] It is the expected value, E(X) = np, with a standart deviation $= \sqrt{np(1-p)}$; n = number of trials, p = the success probability in each trial (Here $n = 47$ and $p = 0.05$).

DISCUSSION

This study helps us draw both general priorities and specific expectations for a particular genre or player category.

We found that a number of heuristics seem to be important for the three types of games. Among the most important of them, many focus on the notion of game variety, rhythm, balance and immersion. These are all aspects that help to maintain the player's attention. The heuristic considered as the most important summarizes well this phrase: "*GamePlay A4: Gameplay is long and enduring and keeps the players' interest*". Other heuristics have also been identified as important. They are dealing with the feedback, the imput/ouput and the consistency of the interface. These heuristics are at the root of any successful interaction and can be found in most current guidelines. On the other hand, some heuristics seem less crucial. They generally focus on specific points of games, such as tutorials, game manual, humor, and the emotional relationship with his avatar. They all are criteria, which are outside the main interaction mode of the game.

In addition to the general recommendations, we found significant differences between game genres that distinctively guide the design of each. For example, the game balance, the interface consistency and the errors minimization are very important aspects in RTS games, while it is rather the environment consistency and the rewards management that prevail in MMORPGs. This shows us that there is more than one way to have fun and this is why the current game industries are so diverse. So diverse that it is not fruitful to base the assessment of all video games with the same universal heuristics. It is interesting to know that other researchers have issued the same warning to film studies, pleading for middle-level research to better account for the specificity of certain kinds of films that were being left out of sight with the prevailing "Grand Theories" [1].

Similarly, we found a large number of heuristics that were judged of minor importance for e-Sports players. A part of them relates to the single player mode, such as the story, the accumulated possessions and AI. Another part relates to "Approachability" criteria [5], such as tutorials, adjustable difficulty levels and basic information required to play. Thus, e-Sports players whose interest focuses on multiplayer and expertise are left unsatisfied. Indeed, when returning the questionnaire, many of them asked why so few criteria were related to the multiplayer mode instead of the single player mode. This is a concrete example that demonstrates the differences of players' expectations.

IMPLICATIONS AND FUTURE WORKS

This study demonstrates that the relevance or importance of a significant number of heuristics differ according to game genres and player types. These findings imply that heuristics should be selected and weighted carefully when evaluating a specific game genre or for evaluating games for specific player types. These findings have implications for game design too because it is impossible (and counterproductive) to optimize a game in all its aspects. Rather, priorities should be identified, according to a certain type of attended game play. In addition, because the player population has greatly expanded, the consideration of specific players' needs, will allow us to find emerging marketing niche to invest.

Because this work has aimed only at providing an overall validation of these statements, substantial in-depth works deserves to be done in order to find relevant criteria for different game genres and players profile. For example, it would be difficult to develop a high-level strategy game if we do not know the specific expectations of the players of these games. To do so, qualitative studies have to be conducted. It should go beyond the over simplified taxonomy used in this study and embrace a larger diversity of game genres and player profiles. In addition, future work should be conducted to investigate the link between heuristics relevance according to game genres, player types and evaluation reliability. Other factors influencing reliability of game evaluation should also be investigated.

Recent scientific studies of video games demonstrate the current appeal of the HCI field in developing theories, methods, and scales to evaluate them. In a broader perspective, we assume that video games studies will also contribute to enrich our general knowledge of user experience theories.

ACKNOWLEDGMENTS

We thank all the volunteers coming from C&C Red Alert 3, Call of Duty 4 and World of Warcraft communities for their time and energy in completing the questionnaire.

REFERENCES

1. Arsenault, D. Video Game Genre, Evolution and Innovation. *Eludamos: Journal for Computer Game Culture. 3 (2009),* 149-176.
2. Bateman, C., & Boon, R. *21st Century Game Design*. Charles River Media, Hingham, Mass, 2006.
3. Caillois, R. *Man, Play and Games* (M. Barash, Trans.). New York: Free Press of Glencoe, 1961.
4. Cockton, G. Woolrych, A., Hornbæk, K. & Frøkjær, E. Inspection-based methods. In A. Sears, and J. A. Jacko (Eds.), *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, Third Edition*, Lawrence Erlbaum Associates, Mahwah, NJ, 2012.
5. Desurvire, H., & Wiberg, C. E. User Experience Design for Inexperienced Gamers: GAP – Game Approachability Principles. In R. Bernhaupt, (Ed.), *Evaluating User Experience in Games*. Springer, London, UK, 2010.
6. Desurvire, H., & Wiberg, C. E. Game Usability Heuristics (PLAY) For Evaluating and Designing Better Games: The Next Iteration. In *Proc. OCSC '2009*, Springer Press (2009), 557-566.
7. Fang, X., Chan, S., Brzezinski, J., & Nair, C. Development of an Instrument to Measure Enjoyment of Computer Game Play. *IJHCI 26*, 9 (2010).
8. Federoff, M. *Heuristics and Usability Guidelines for the Creation and Evaluation of Fun in Video Games*. Master of Science, Indiana University, Bloomington, IN, 2002.
9. Jørgensen, K. Players as Coresearchers: Expert Player Perspective as an Aid to Understanding Games. *Simul. Gaming* 43 (2012), 374-390.
10. Koeffel, C., Hochleitner, W., Leitner, J., Haller, M., Geven, A., & Tscheligi, M. Using Heuristics to Evaluate the Overall User Experience of Video Games and Advanced Interaction Games. In R. Bernhaupt (Ed.), *Evaluating User Experience in Games*. Springer, London, UK, 2010.
11. Lazzaro, N. *Why we play games: Four keys to more emotion without story*. (Report). XEO Design Inc, 2004.
12. Livingston, I. J., Mandryk, R. L., & Stanley, K. G. Critic-proofing: how using critic reviews and game genres can refine heuristic evaluations. In *Proc. Futureplay '10*. ACM Press (2010), 48-55.
13. Nielsen J. (1994). Enhancing the explanatory power of usability heuristics, In *Proc. CHI 1994*, ACM Press (2004), 152-158.
14. Pinelle, D., Wong, N., & Stach, T. Heuristic evaluation for games: usability principles for video game design. In *Proc. CHI 2008*, ACM Press (2008), 1453-1462.
15. Pinelle, D., Wong, N., & Stach, T. Using genres to customize usability evaluations of video games. In *Proc. Future Play '08*. ACM Press (2008), 129-136.
16. Scapin, D. L., & Bastien, J. M. C. Ergonomic criteria for evaluating the ergonomic quality of interactive systems. *Behaviour & Information Technology, 16* (1997), 220-231.
17. Sweetser, P., & Wyeth, P. GameFlow: a model for evaluating player enjoyment in games. *Comput. Entertain 3*,3 (2005).
18. Yee, N., Ducheneaut, N. & Nelson, N. 2012. Online gaming motivations scale: development and validation. In *Proc. CHI 2012*, ACM Press (2012), 2803-2