Does Usability affect Engagement in Long-Term Game Experiences?

Alessandro Febretti
Milestone Games
Via Gustavo Fara 35 - 20124 Milano, Italy
and
University of Illinois at Chicago
Electronic Visualization Laboratory
842 W. Taylor Street - 60607 Chicago (IL), US
Alessandro.febretti@gmail.com

Franca Garzotto
Politecnico di Milano
School of Industrial Design and
Department of Electronics and Information
Piazza Leonardo da Vinci 32- 20133 Milano, Italy
franca.garzotto@polimi.it

Abstract
We investigate how interface usability influences user engagement in long-term computer games, and assess the relevance of usability versus playability factors on the overall experience engagement.

Keywords
Interactive game, engagement, usability, playability

ACM Classification Keywords
H5.2. User Interfaces – *Evaluation/Methodology*

Introduction
Many existing computer games developed at industrial level represent a category of applications that can easily be associated to a long-term interaction experience. Gameplay sessions normally last one hour or more, and designing games for a total play time in the order of tenths of hours is a common industry practice. Several games (especially the ones based on multiuser interaction, like MMOGs - Massive Multiplayer Online Games [3]) can even offer a theoretically unlimited gameplay experience.

Games, like any other computer application, expose an interface between the user and the application itself (in this case, the game world). Since interaction is mediated through the game interface, usability is, in principle, important for the quality of the user experience during playing. In particular, we may expect

that usability problems could negatively affect user engagement in the game.

On the other hand, a games' primary objective is not being efficient and effective in completing a specific task: they instead need to succeed in providing *overall fun* to the user, which is the fundamental factor to engage users and motivate them to continue use of the game over time.

Therefore, the question we want to address is the following: when considering the entire long-term interaction experience, does *usability truly affects engagement*, or are other *factors* (for instance, playability features related to *enjoyment*) that may overwhelm the influence of usability defects?

In this paper we report a study in which we evaluated a number of usability and playability factors in a group of eight commercial long term games, empirically measured the user's overall engagement, and investigated the correlation between engagement and those factors. For our study, we adopted both inspection methods and empirical methods, involving a team of twenty game design experts as inspectors of usability and playability, and 47 users for the empirical measurement of engagement.

Heuristic Class	Heuristics
Concentration &	7
Challenge	2
Player Ability	4
Control	1
Objectives &	4
Social Interaction	2

Table 1. Categorization of gameplay heuristics

Heuristic Class	Heuristics
Customization	3
Input	4
Game Views	2
Interface	4
Game Menu	1

Table 2. Categorization of usability heuristics

Usability and Playability Heuristics

We conducted our analysis of usability and playability factors in the targeted games using a heuristics-based method. This approach has a number of advantages over user based evaluation, mainly because of lower operative cost and higher effectiveness in detecting specific and fine-grained design defects [10].

From the various proposals of playability heuristics available in the current state of the art [1,2,5,12], we selected twenty-two heuristics we deemed more significant for our games, and we categorized them in seven different *classes* (Table 1). Similarly, the general usability heuristics found in literature [10], were filtered to discard those that were not significant in a computer game scenario, we slightly customized some of them, and introduced some new ones basing on industry practice [4, 5, 11]. Five heuristics classes were identified (Table 2) for a total of fourteen usability heuristics. These sets of categorized heuristics (which will not be described here due to limited space) were used as a common and consistent schema among all evaluators during the inspection process.

Evaluating Engagement

A qualitative measure of user engagement has been first estimated through supervised and unsupervised user testing. Each game was submitted to groups of five to eight users. Each user had to complete a number of scenarios that were custom for each game and covered the most significative gameplay aspects of the game being analyzed. Evaluators' observations during gameplay, interviews and questionnaires at the end of the play experience provided the qualitative and quantitative data that allowed us to elaborate an initial assessment of the engagement level offered by the game. These measures were then compared and integrated with data coming from game rating aggregators available on specialized web sites for game quality assessment [7, 9], to gain a more precise and reliable measure of game engagement.

Investigating Correlations

Heuristics-based inspection of each game resulted in a set of measures for each usability and playability attributes. Through a weighted average of these values we obtained a global estimation of the game compliance to the chosen usability and playability criteria. Statistical correlation analysis was carried out using Spearman's rank correlation test [8] between categorized and global indicators, and the game engagement rating. The resulting correlation coefficients between global playability/usability estimates and game engagement are shown in Table 3.

Correlation type	Correlation value
Playability ↔ Engagement	0.760 (> 0.738, p=0.05)
Usability ↔ Engagement	0.515

Table 3. Correlation between global heuristics / usability estimates and game engagement

Comparing the estimated correlation coefficients with thresholds for 0.05 significance level allows us to identify a *significative relationship* between the analyzed *gameplay heuristics* and the actual level of user *engagement* in the game. The *relationship* between the considered *usability* factors and *engagement* instead, does *not* appear to be *statistically relevant.*

This difference in correlation levels is evident also in estimations between single heuristics classes and game engagement (figure 1).

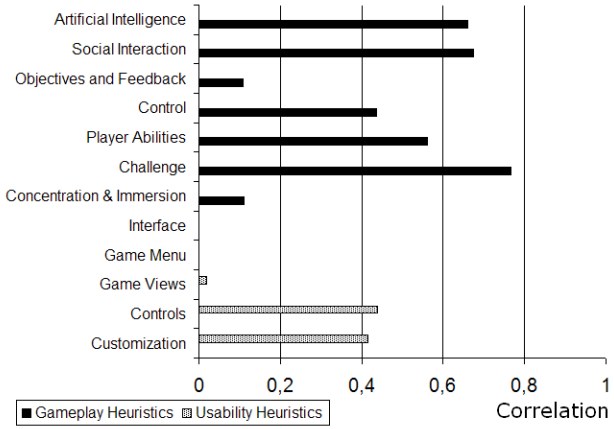

Figure 1. Correlation between heuristics classes and game engagement

Correlation levels for gameplay heuristics classes are significantly higher on average than correlations for usability classes. It is worth noting how heuristics classified in the *challenge* category appear to be strongly linked with game engagement. This class of heuristics cover the aspect of *balance* in the game challenge level, in relation to both player initial skills and progress throughout the game. The fundamental nature of these aspects in gameplay design is well known [1, 2, 12], this result is therefore not surprising. It is also interesting to observe how the only two usability classes that show some relationship with engagement are the ones that deal with the *dynamics of user interaction*: *user activity* with the interface has a more effect on engagement than simple exposure to the interface.

Discussion and Conclusions

Although a number of studies have investigated separately usability and playability in gaming experiences [1, 2, 4, 5, 6, 11, 12], to our knowledge this is the first attempt that explicitly investigates the relationships among these factors and engagement, in the context of long term professional games.

Our study shows that engagement appears to be significantly related to playability factors in long-term professional computer game experiences, but statistical analysis highlights a *low correlation between usability and engagement* in the same context. The latter result is apparently surprising, since during user testing we observed a decrease in user engagement in the game experience when a usability problem manifested itself during a play session.

To interpret this phenomenon, we can consider the concept of *Density of Usability Defects,* or *DUD,* defined as $DUD = UP/T$ where *UP* is the number of usability problems for the game experience duration T. In other words, the density indicates the amount of usability defects a user will face for a specific gameplay length.

In long term commercial computer games like the ones considered in our work DUD is low, i.e., usability problems have an *episodic, sporadic* nature, and a specific interface or interaction defect manifests itself in a limited number of hot spots inside the gameplay experience. That is because an industrial game is subject to strong quality assessment (QA), and a frequent user interface problem would probably be noticed, and removed, during the QA phase of industrial production.

If the Density of Usability Defects is low and a specific usability problem manifests itself during a play session, the decrease in user engagement in the game experience seems to be *temporary* and *local*, remaining confined to the moment of defective interaction without significantly affecting the *overall* engagement. On the long-term, factors that have a *constant* influence on the experience (like the ones captured by playbility heuristics) appear to be more relevant for the *complexive engagement* of the game.

The results of our study may not apply to games with shorter play sessions, like web-based minigames or casual games. A game of this kind having the same usability problems as a long-term game would tend to manifest these defects more: in other words, usability defect density could increase due to the simpler and shorter gameplay structure. In this case, we may expect that usability can have a deep influence on overall engagement. Similar considerations apply to games developed by non-professional teams, or games that do not go through a consistent quality assessment phase. In this scenario, even complex and long gameplay structures can be damaged by severe, unidentified usability problems that influence the entire play experience.

Verifying these methodological hypotheses is the subject of our future work.

References

[1] J. Chen, “Flow in Games”, http://jenovachen.com/flowingames/thesis.htm

[2] Desurvire, H., Caplan, M., Toth J.A., “Using Heuristics to Evaluate the Playability of Games”. Proc. CHI 2004, April 2004, pp.457-466

[3] Ducheneaut, N., Yee, N., Nickel, E., Moore, R., “Alone Together” Exploring the social dynamics of massively multiplayer online games”. Proc. CHI2006, pp. 407-416

[4] Federoff, M. “User Testing for Games: Getting Better Data Earlier”. Game Developer Magazine, June 2003

[5] Federoff, M. “Heuristics and Usability Guidelines for the Evaluation of Fun in Video Games”. Thesis at the University Graduate School of Indiana Univ., Dec. 2002

[6] Garzotto F. “Investigating the educational effectiveness of multiplayer online games for children” Proc. ACM IDC 2009, pp. 29 – 36

[7] “IGN”, IGN Entertainment, http://www.ign.com

[8] Graziano A.M., Raulin M.L. “Research Methods – A process of Inquiry”. Pearson, 2007

[9] “Metacritic”, http://www.metacritic.com

[10] Nielsen, J., Molich, R.L., “Usability Inspection Methods”, John Wiley & Sons, 1994

[11] Pinelle D., Wong N., Stach T., “Heuristic Evaluation for Games: Usability Principles for Video Game Design”, Proc. CHI 2008, April 2008, pp. 1753-1762

[12] Sweetster P., Wyeth P., “GameFlow: A Model for Evaluating Player Enjoyment in Games ”, ACM Computer and Entertainment, 3, 3. 2005, pp. 1-23.