

Quality Assessment of Digital Games

Thiago Alves, Luiz de Mello

Pontifícia Universidade Católica de Minas Gerais, Brazil

Abstract

The need for this work was perceived while examining the methodologies used to evaluate the quality of modern and past digital games or, rather, lack thereof. This work is therefore an initiative to collect data about the necessity or not of establishing a methodology to evaluate the area. For this purpose we designed a checklist to be completed by players and specialists to collect data about its quality.

Keywords: Quality of digital games, Quality Evaluation, Digital Games.

Authors' contact:

thiagoaalves@gmail.com; luizdemello@ymail.com

1. Introduction

The market in the industry of Digital Games has become extremely competitive due to the advent of new technologies together with the new generation of consoles like Playstation 3, Xbox 360 and Nintendo Wii, as well as development for portable devices, such as high performance mobile cellphones (e.g. iPhone), PSP and Nintendo DS. All of that has made many companies invest in the development of digital games.

But only investment in development does not guarantee success of the games that are released; you must ensure that the game being developed is interesting and can attract the attention of players and still be fun for them. But how can we know if a game is good and fun? How can we measure the quality of games not only from a technical point of view (e.g. quality of the software), but also from a conceptual point of view (e.g. the playing experience)? A problem that is happening now in this industry is that there is no formalized material specifically about the quality of digital games. What exists is the adoption of a set of best practices that are common knowledge within the industry. Because of this problem, the contribution of this work is of high value, because there is little solid information which credibility is verifiable and those found do not always meet the expectations.

It is noteworthy, however, that the approach taken in this work was mostly classical, that is focused on projects based on a single area. What was observed is that the sub-areas could not be simply grouped as one (as the final product sometimes appear to be, 'just' a package). Its specific sub-areas (software development, visual arts, music and sounds, etc.) sometimes can retain a level of independence that made it unwise to group. Yet, it was also observed that

at the same time those areas can display a high level of interdependency in different moments (different games or even different phases of the production). That is enough to make a complete division of each area also unwise. For example, a game's soundtrack may consist of songs considered by musical standards only average, but that were perfect for specific moments of the game. More elaborate and musically appealing songs would be a poorer choice since they would produce undesired effects, like distracting the player or making the voices of NPCs hard to hear in a very critical action scene. The best approach needs to be hybrid and adaptative. It needs to be able not only to consider the individual aspects of each subfield, but also their integration and also be adaptable to each game. This may sound like a drawback on this work, but it actually just reinforces the conclusion: there are no formal methodologies to assess the quality of digital games that are sufficiently satisfactory.

2. Related Work

During the development of this paper, few related works were found, what made us believe that this work is indeed an important initiative. The article Evaluation of Quality in Computer Games [Sandvik, K. 2006] is a similar work, although its focus differs greatly. It touches the lack of proper standard means to evaluate the quality of games, but it has emphasis on the social impact of games. Being able to properly judge the quality of a game is one step to evaluate their impacts, while this work focuses on the process of evaluating the quality of games itself and proposes one standard system.

3. General Recommendations

This paper aims to propose a model to assess the quality of digital games because currently there is no particular pattern for this type of product. The ultimate objective is to propose a model that can audit the quality of digital games both new and old. The standard will be based on the ISO / IEC 9126:1991. This standard presents a set of six main features and each feature has a major sub-set of features. Not all features and sub-features are applicable to the context of digital games. Adjusting those and creating a new feature and its sub-features is where this work comes into prominence. The additional feature will be called "Game" and has its own set of sub-characteristics.

3.1 Quality model Specification

This work is based on the quality model from ISO / IEC 9126, this model features a set of six features and each feature has a set of sub-features. Not all features and sub-characteristics fit the concept of digital games for both this work and proposed a new feature called “Game”, which had its own set of sub-features that aims to assist in the evaluation of digital games. The ISO / IEC 9126 introduces the following set of features and their sub-characteristics: Functionality (Suitability, Accuracy, Interoperability, Security Access and Compliance), Reliability (Maturity, Fault Tolerance and Recoverability), Usability (Intelligibility, Operational and Apprehensibility), Efficiency (Behavior in Time and Behavior in resources), Maintainability (analyzability, modifiability, stability and testability) and Portability (adaptability, capacity to be installed, the Standards Compliance Portability and Ability to Replace).

Some of these characteristics and sub-features are not entirely applicable to digital games. With this in mind, we will propose the creation of a new feature that will present its own set of sub-features that aim to cater exclusively to digital games. This feature is called Game, and its sub-features are: Gameplay, Fun, Graphics, Story, Music and Originality.

4. Procedure for Evaluation and Classification

The ultimate goal of this study was to propose a model to evaluate the quality of games in the form of checklist along with a classification model.

4.1 Metrics

To evaluate and classify the quality of digital games 61 metrics were defined based on the standards ISO / IEC 9126 and the features developed specifically for digital games.

Functionality

Adequacy

1. How do you rate the grouping of menus?
2. How do you rate access to the menus in the game?

Accuracy

1. How do you rate the response of the controls on the screen? Do they give the impression of being late or running at the wrong time?
2. How do you rate the absence of errors in control? When a certain command is inserted is corresponding action triggered and executed correctly?

Interoperability

1. How do you rate the options on the screen?
2. How do you rate the need for these options? What really is essential is presented first or is it needed to navigate through menus to access those options?

Conformity

1. How do you rate the game controls compared with the controls of the games of the same genre of previous installments of the franchise?
2. How do you rate the playability of the game compared to the controls of the games of the same genre of previous installments of the franchise?

Security

1. How do you rate the form used to save the game? Do you consider it safe?
2. How do you rate the safety of commands? Is it possible that using certain combinations of commands during the game it triggers an error?

Reliability

Maturity

1. How do you rate the absence of errors in the game?
2. If any mistake happens, how do you rate the handling of this error in the game?

Fault Tolerance

1. How do you rate the fault tolerance within the game? When a mistake happens, can the game continue to run normally?
2. How do you rate the ability to recover data when an error occurs while running the game? Does the game allow an immediate save before closing when a fatal error occurs or allows the player to recover progress not manually saved when restarted after a crash?

Recoverability

1. How do you rate the absence of errors during a game save? Does it show no problems on purpose (i.e. poor design of the process) during a save?
2. How do you rate the recovery of saves? If there is any problem during an override is it possible to recover the previous version easily?

Usability

Intelligibility

1. How do you rate the tutorial starting at the beginning of the game? Is the tutorial easy to understand or too complex and difficult?
2. How do you rate the presentation of controls within the game? Are all commands presented in the game or is it necessary to consult or search sources outside of the game?

Apprehensibility

1. How do you rate the learning process within the game? Is it natural and easy to learn how to learn the game?
2. How do you rate the information presented during the game? Is it easy to assimilate them or is it complex and difficult to understand what is being reported.

Operational

1. How do you classify the commands of execution of the game?
2. How do you rate the performance of joint command during the execution of the game? Is it easy to

intentionally run more than one command simultaneously and are they executed as expected?

Efficiency

Behavior in Time

1. How do you rate the time used for loads in the game?
2. How do you rate the frequency of loads in the game?

Resources' Behavior

1. How do you rate the usage of resources of the platform where the game is running? Do you think the game is using the full potential of the platform or does it manage them poorly?

Maintainability

Note: Initially neither this feature and its sub-features were strongly marked as applicable to digital games, being excluded from the evaluation process. After a few conversations with specialists and additional research it was possible to create two metrics related to the concept of maintainability within digital games.

1. How do you rate the launch of paths for the correction of errors?
2. How do you rate the frequency in launching paths for correcting errors? If you found an error and takes up much to get a fix for that error.

Portability

Note: This feature had to be contextualized to digital games, as its original sub-features were not precisely applicable to digital games.

1. The game features versions for other platforms and environments?
2. These versions have the same standards?
3. Is there any communication between the game and its other platforms?
4. The game has a feature so it can be used on other platforms in their entirety or partially?

Game

Gameplay

1. How do you rate the interaction of menus within the game?
2. How do you rate the navigation of menus within the game?
3. How do you rate the game controls?
4. How do you rate the response time for the game controls?
5. How do you rate the ease of understanding the commands?
6. How do you rate the handling of the camera?
7. How do you rate the layout of the camera in the scene?

Fun

1. How do you rate the difficulty that the game offers?
2. How do you rate the riddles and puzzles found throughout the game?
3. How do you rate the entertainment offered by the game itself?

Graphics

1. How do you rate the quality of artistic objects and characters that exist within the game?
2. How do you rate the level of detail of objects and characters that exist within the game?
3. How do you rate the artistic quality of the scenarios?
4. How do you rate the level of detail of the scenarios?
5. How do you rate the vision of objects belonging to the scenario that has interaction with the player?
6. How would you classify the objects belonging to the scenario that has interaction with the player?

Plot

1. How do you rate the story in a general context of the game?
2. How do you rate the main characters existing within the game?
3. How do you rate the secondary characters existing within the game?
4. How do you rate the chronological facts of history in the game?
5. How do you rate the context or environment and the world that the game and set?
6. How would you classify the course of events and happenings within the game?
7. How do you rate the central plot found in the game?
8. How do you rate the end of the game?

Music

1. How do you rate how remarkable was the soundtrack of the game?
2. How do you rate the integration of the soundtrack to the game?
3. How do you rate the diversity of the soundtrack during the game?
4. How do you rate the level of cohesion of the music of the game in its entirety?
5. How do you rate the relationship of music with the history of the game?

Originality

1. How do you rate the central idea of the game, even though similar to some other game does it still give the impression of being unique?
2. How do you rate the new features that can be explored within the game, even if similar to other games resources does the whole set of them seem to be unique?
3. How do you rate the story and the game world, even though similar to some already played does it passes the feeling of being unique?

4.2 Score levels

Each metric may receive one of the following responses and each option receives a specific number of points that will compose the final score:

Functionality

Responses have the following weights: Very Good = 4, Good = 3, Fair = 2, Poor = 1, Not Applicable = 0.

Reliability

Responses have the following weights: Very Good = 4, Good = 3, Fair = 2, Poor = 1, Not Applicable = 0.

Usability

Responses have the following weights: Very Good = 4, Good = 3, Fair = 2, Poor = 1, Not Applicable = 0.

Efficiency

Responses have the following weights: Very Good = 4, Good = 3, Fair = 2, Poor = 1, Not Applicable = 0.

Maintainability

Responses have the following weights: Very Good = 4, Good = 3, Fair = 2, Poor = 1, Not Applicable = 0.

Portability

Responses have the following weights: Yes = 4, No = 0.

Game

Responses have the following weights: Very Good = 4, Good = 3, Fair = 2, Poor = 1, Not Applicable = 0.

4.3 Review

According to the checklist, the score will show values that will meet the range of 0 points and 244 points. The minimum value considered for the score is 0 points which is obtained by selecting 'Not Applicable' or 'No' on all metrics of the checklist. The maximum score is 244 points which is obtained by checking the 'Very Good' or 'Yes' option in all metrics of the checklist.

To classify the scores it was decided to use an ordinal scale. According to Teixeira [2005]: 'The ordinal scale is a measuring instrument that sorts the objects according to the degree that they possess a given attribute.' This model seemed to be the most appropriate as it allows a less rigid and more forgiving final classification, one that better accommodates the particularities of digital games.

The final classification of a game occurs accordingly to the following table:

Rating	Score
A Poor Game	From 0 up to 100 points
An Average Game	Above 100 up to 150 points
A Good Game	Above 150 up to 200 points
An Excellent Game	Over 200 up to 244 points

Table 1: Rating

A Poor Game: It has few interesting aspects. This kind of game will more likely not be successful. Gamers will probably get tired quickly with no real replay value, if it is worth completing even once.

An Average Game: It is not very interesting, but better than the poor games. It can keep the players

entertained for a while, but not for long. It is a decent pass time, but nothing memorable.

A Good Game: It is interesting and keeps players entertained for a long time. Players will want complete the game to explore its various interesting features. They may not be considered the best of their kind, but will give the player a feeling that got a decent value in return for the money they invested.

An Excellent Game: It can be considered an "Epic". These games are those that will be played over and over and that players will often find themselves playing again after a long break. The game is so strong that he will hardly be forgotten as time passes.

Conclusion

A game cannot be justly classified only by a single person and needs a set of people for that, because it is important to find an 'average opinion', rather than just one based on personal preferences. But for that to happen there must be a proper way so that everyone can opine in an uniform manner that makes it simple to consider and combine those opinions into a single final verdict. The author believes that this work can allow that to happen because it considers both classic elements already employed to evaluate elements of the more mature software industry as well elements specifically designed to cater to the game industry.

The overall importance of this work is to help the professionals in the game industry to have a proper formal criteria to define if the game being produce is already good or what it is lacking to be.

Acknowledgments

We would like to give thank to: God; [Professor's name] for all the guidance, motivation and encouragement given to us, directly and indirectly; [musician's name], 8th Period Course Bachelor of Viola at the Music School of UFMG, for his explanation of the role of music in the games; and to light of Kilitias that warms the hearts of those who believe.

References

- ISO/IEC 9126: Software Engineering - Product quality - Part 1: Quality model (1991) Available from: ftp://www.puc-campinas.edu.br/pub/professores/ceatec/tobar/tcc/info_is_oiec9126-1.pdf [Accessed 15 April 2010].
- SANDVIK, K. 2006; Evaluation of Quality in Computer Games, *Nordicom Review* 27(2): 269-284.
- TEIXEIRA, G. 2005; Conheça o que São Escalas de Medida [online] Ser Professor Universitário. Available from: <http://www.serprofessoruniversitario.pro.br/ler.php?modulo=21&texto=1304> [Accessed: 01 July 2010].