Usability Evaluation for Computer Games

Introduction

This handout examines the practice of usability evaluation as it applies to computer games. Though it may surprise you to hear it, evaluating games is not that different from evaluating any other piece of software. In games, as with other software, evaluation is a continuous process which occurs throughout product development as part of the iterative development lifecycle. In fact, most of the methods you have already learned can be adapted for use in game evaluation.

What Makes Games Different

Though game evaluation may employ familiar methodologies, make no mistake – games are different from other software.

Think of the applications you use every day – software like Mozilla Firefox, Microsoft Word, and Adobe Photoshop. Can you see the common theme? Each of these products facilitate an existing task, whether it's surfing the net (Firefox) writing a letter (Word) or editing an image (Photoshop). To put it another way, each of these products are validated by external goals – people like Word because it helps them to perform something they already wanted to do.

Games are different. Because a game does not provide any external reward, they must be intrinsically rewarding. That is to say, using the software must be a satisfying experience in and of itself. That is the key difference between games and applications, and if you only take away one thing from this document, I want it to be this:

A good application allows its users to *do* something.

A good game makes its users *feel* something.

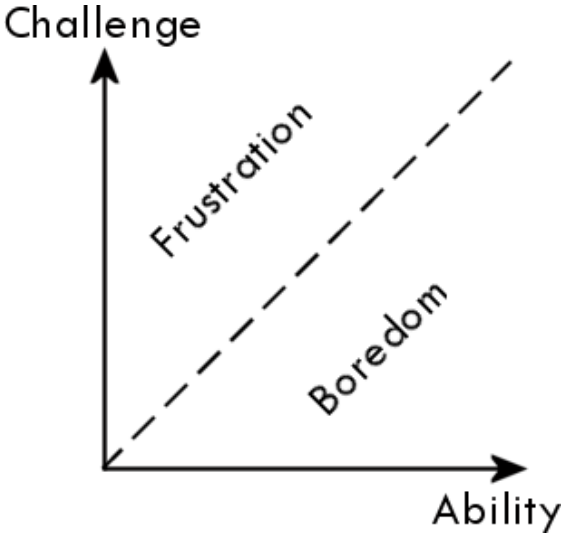
But which feelings should a good game evoke? Ostensibly, games are designed to provide amusement, so a good game should create the feelings associated with amusement – happiness, excitement, and so on. The feelings of "fun" that a game creates can (roughly) be broken up into two concepts: hard fun and easy fun [5].

Hard Fun

Hard fun describes the rewarding sensations that a player experiences when he or she overcomes a challenge. Hard fun is closely linked to learning [3]; when a player overcomes a challenge it indicates that he has developed and employed new skills to overcome a formerly difficult task.

Hard fun can only exist in a context where the challenge is slightly above the player's existing abilities. Since players are continually learning as they play, a good game must increase its challenge at a comparable rate in order to provide a continual stream of difficulties to overcome, and consequently, a continual stream of hard fun.

The process of balancing game difficulty with player ability is perhaps, the most significant challenge faced by game developers. If a game's challenge increases faster than a player's ability, that player will become frustrated and likely quit playing. Contrariwise, if the challenge increases too slowly, players will become bored and under-stimulated. As a result, a large part of game evaluation is dedicated to identifying and eliminating frustration and/or boredom.



The ideal difficulty curve for games - as the player's ability increases, so does the challenge, creating a continual supply of "hard fun".

Testing for and Detecting Hard Fun

Creating hard fun is all about maintaining the balance between challenge and ability. Since players' innate abilities will naturally vary across demographics, it is important to balance your game according to the needs of your target audience.

It is nearly impossible to achieve this balance without hands-on testing ("playtesting") with players from your target audience. By having participants play the game as they normally would, and observing their behavior, body language, and verbal outbursts, you can learn a great deal about your game, and how your players are experiencing it. You may find that you've overestimated players' learning abilities, and accelerated the challenge too quickly. On the other hand, you may have made the game too easy – providing gratuitous hints when the players would prefer to puzzle out a solution themselves.

There is any number of player behaviors which can indicate a breakdown in your game's "hard fun" – in order to pick up on them, it's necessary to be a little intuitive, and a little empathic. For your reference, here are a couple situations you may observe during testing, and what they may mean:

- **The player is repeatedly failing a task:** For a player, this can be very frustrating. It may indicate that the player does not know what to do, or that the player has not mastered the skills which are crucial for success. You may observe this behavior directly, or you may hear the player say things like: "This is too hard", "I can't figure this out" or "How do I do this?"

  You can address this problem by providing in-game hints, reducing the complexity of the task, or giving the player the opportunity to practice his skills in a consequence-free area before the big task.

- **The player is proceeding through the game without any problems at all:** The opposite of the above problem. Although this may *seem* like a good thing, it is likely not enjoyable for the player, since it suggests that he is not applying his

skills to the utmost. You may hear the player say "This is too easy."

You can address this by increasing the difficulty of the game as a whole, or by allowing creating (optional) additional challenges which extra-skilled players can use to test themselves.

- **The player is lost, or does not know how to proceed:** You may witness the player backtracking, wandering in circles, and trying the same things over and over again. This can be very frustrating for players. You may hear the player say "I'm lost", "I don't know where to go" or "What do I do now?"

  Once again, this can be addressed through hints and other cues. The use of cues is one of the strongest skills in a game designer's arsenal. For example, you may draw attention to a certain door, by having a ceiling light flickering over top of it; this will naturally compel the player to investigate. Although you drew the player's attention there, it was done in such a subtle way that the player will still feel like it is he is acting of his own accord when he goes to investigate.

- **The player is repeatedly using the same strategy/tool/weapon:** This could be indicative of boredom. It suggests that the player has found one strategy which is applicable in nearly all cases – and thus, he does not need to think, or adapt to new circumstances. In short, it suggests that he is done learning.

  You can address this by lowering the effectiveness of the chosen strategy, or by introducing new challenges against which the chosen strategy will prove ineffective.

- **The player is repeatedly performing the same or similar tasks:** This is a surefire way to bore your players – spending too much time working on the same problem over and over again is not mentally stimulating.

  This likely reflects a problem with your game's design. Break up long, tedious sections, and give players a chance to "cool down" between difficult tasks. By varying the challenges that the player faces, you can hold their interest for much longer.

- **The player is yelling expletives/banging his head on the keyboard/etc:** This really happens, on occasion. After all, gaming is an emotional experience. This is a pretty sure sign that *something* is wrong. If you observe this behavior, talk to the player, and see if you can identify the source of his frustration.

In general, the signs may not be this explicit; a player who sits stone-faced through an entire playtest may still provide valuable insight for your design. For this reason, other testing methods are commonly used in conjunction with playtests to increase the amount of available data. These include: having players think-aloud while they play, administering questionnaires to playtesters either during or after the test, and conducting oral interviews (Thompson). Data-logging (including video-taping) your

playtests can also yield valuable results; by tracking statistics such as where your players are spending their time, where they are dying and what items and tools they're using, you can identify possible sources of frustration and other problems in your design [2, 8].

Easy Fun

Whereas hard fun requires directed action on the part of the player, easy fun is almost entirely passive. Easy fun describes the pleasurable sensations of pure experience and exploration – the joy of interacting with compelling characters, of hearing a beautiful piece of music, of witnessing a plot revelation, and of feeling as though you are a part of a fully-realized world. Easy fun is akin to the joy a person receives through traditional, non-interactive media such as movies, television, and music.

Testing for and Detecting Easy Fun

Easy fun is quite subjective – what may be very "cool" to one player may seem clichéd to another. Therefore, with easy fun (as with hard fun) hands-on playtesting will help you to determine what your audience will enjoy. Here are some tell-tale signs that you can use to determine if your players are experiencing easy fun:

- **The player is involuntarily emoting:** You may witness players smiling, squealing in terror, or sitting stunned, with mouth agape. Whatever the case, these involuntary physical reactions suggests that the player is experiencing a vivid emotional state. These reactions may be accompanied by emotional vocalizations such as "Wooow…", "Ohmigod!" or "Yes!"

  When you see this happening, it's a good idea to ask the player what prompted this – especially because they may not even realize they're doing it! You'll want to make sure they're reacting to the right things. If your players are laughing at a scene which is supposed to be scary, you're probably doing something wrong.

- **The player is taking time to explore, experiment or admire:** This can manifest itself in many different ways; any time you witness the player engaging in undirected exploration, they can be said to be engaging in easy fun.

After your playtesters have finished, questionnaires and interviews can help you to understand their experience. Players who have enjoyed a game will be eager to talk about their experience, so asking open-ended questions about the player's emotional highs and lows (for example, "What was your favorite part of this game?" or "Having completed this playtest, which experience sticks in your mind?") will help you to understand where your product succeeds and where it fails.

Focus groups - particularly in the early stages of development – can also help developers to create easy fun, as they can provide general insights into the likes (and dislikes) of your target audience [1]. These insights can then be integrated into your game's design from the very beginning.

Physiological Feedback Evaluation

Playing a game is emotional experience, and in the course of gameplay, players will often manifest their emotions physically, by grimacing, smiling or even gawping in open-mouthed awe. Some researchers have suggested that this phenomenon could be leveraged as a method to evaluate games [6]. By monitoring the physiological state of a player (including their heart-rate, the conductivity of their skin, and the configuration of their facial muscles) and matching these values to the known parameters for certain emotional states, it would be (theoretically) possible to determine the emotional state of the player as they play. This data could then be used to justify the success (or lack thereof) of the game. For example, if the playtesters spent a large portion of their time smiling, then one might infer that the game is amusing. To the best of my knowledge, this method is not actively employed by any professional game developers. I present it merely for your consideration.

Conclusion

Since you, the reader, are already experienced in human-computer interaction and/or human factors, this I have selected to focus this handout on that which you do not already know: the ways in which evaluating games differs from a typical usability evaluation.

Games are not motivated by an external goal – if the game is good, then the very act of playing should reward the player with positive feelings. Taken together, these "positive feelings" are referred to as the "fun" of the game, and can be roughly divided into two categories: hard fun, and easy fun. Hard fun is the enjoyment a player receives from learning, devising strategies, and applying these strategies to overcome problems. Easy fun is the passive enjoyment received when the player experiences something personally moving within the game, such as an interaction with an amusing character or a spectacular and satisfying explosion.

Usability evaluation for games is a broad topic – far too broad for me to do justice within the scope of this handout. If this has interested you at all, I invite you to research further, beginning with the bibliography, below.

Bibliography

[1] Fullton, B. "Beyond Psychological Theory: Getting Data that Improve Games" *in the Proceedings of the Game Developer's Conference 2002*, CMP Media. (2002) http://www.gamasutra.com/gdc2002/features/fulton/fulton_01.htm

A very useful paper, published in association with Microsoft Game Studios. This paper outlines the criteria which usable feedback must have. It then examines some methods which are popularly used, and how they stack up against these criteria. Although this paper was written with an emphasis on games, it's general enough that it could be used for almost an evaluation process.

[2] DeRosa, P. "Tracking Player Feedback To Improve Game Design" http://www.gamasutra.com/view/feature/1546/tracking_player_feedback_to_.php?page=1

A short, but interesting article. They show some potential uses of data-logging here.

[3] Koster, R. "A theory of fun for game design" Scottsdale, AZ: Paraglyph. (2004)
Although this book isn't strictly about game evaluation, it does address what it means for a game to be "fun" – a crucial topic for anyone who hopes to evaluate games.

[4] Laitinen, S. "Better Games Through Usability Evaluation and Testing"
http://www.gamasutra.com/features/20050623/laitinen_01.shtml
Accessed on October 24th, 2007.

An excellent article which gives a concrete example of how usability testing in games is done. This article documents the testing process for the game Shadowgrounds, which included both an "expert evaluation" as well as a user test. The article shows real usability problems, along with their proposed solutions and the developer's comments.

[5] Lazzaro, N. "Why we play games: Four keys to more emotion without story" Technical report, XEO Design Inc. (2004)
http://www.xeodesign.com/xeodesign_whyweplaygames.pdf
Accessed on October 24th, 2007.

This is the document which introduced me to the concepts of "hard fun" and "easy fun". This paper presents two additional concepts which they feel contribute to a player's enjoyment.

[6] Mandryk, R. L. and Atkins, M. S. "A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies." *Int. J. Hum.-Comput. Stud.* 65, 4 (Apr. 2007), 329-347
http://www.reganmandryk.com/pubs/mandryk_ijhcs2007.pdf
Accessed on October 24th, 2007.

This paper suggests a method by which one could use physiological feedback to measure the emotion of players.

[7] Microsoft Game Studios. "MGS User Research – Research and Publications"
http://mgsuserresearch.com/publications/
Accessed on October 24th, 2007.

The homepage for Microsoft Games Studio's User Research Group. This is *the* definitive source for game evaluation information on the internet. Here, you can download many insightful papers and presentation dating back to 2002.

[8] Thompson, Clive. "Halo 3: How Microsoft Labs Invented a New Science of Play." *Wired.* (Aug. 2007)
http://www.wired.com/gaming/virtualworlds/magazine/15-09/ff_halo?currentPage=1
Accessed on October 24th, 2007.

This Wired article details Bungie Software's experience in evaluating Halo 3. It gives you a good idea of the type of things game evaluators have to look out for, and some of the techniques they use. Halo 3 sold over $300 million in its first week, so they must have done something right.