

THE CRITICAL EFFECT: EVALUATING THE EFFECTS AND USE OF VIDEO GAME REVIEWS

A Thesis Submitted to the
College of Graduate Studies and Research
in Partial Fulfilment of the Requirements
for the degree of Master of Science
in the Department of Computer Science
University of Saskatchewan
Saskatoon

By
Ian Livingston

©Ian Livingston, June 2011 All rights reserved

PERMISSION TO USE

In presenting this thesis/dissertation in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis/dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis/dissertation work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis/dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis/dissertation.

Requests for permission to copy or to make other uses of materials in this thesis/dissertation in whole or part should be addressed to:

Head of the Department of Computer Science

176 Thorvaldson Building

110 Science Place

University of Saskatchewan

Saskatoon, Saskatchewan S7N 5C9

Canada

DISCLAIMER

Reference in a thesis/dissertation to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement, recommendation, or favouring by the University of Saskatchewan. The views and opinions of the author do not state or reflect those of the University of Saskatchewan, and shall not be used for advertising or product endorsement purposes.

ABSTRACT

Game reviews play an important role in both the culture and business of games – the words of a reviewer can have an influential effect on the commercial success of a video game. While reviews are currently used by game developers to aid in important decisions such as project financing and employee bonuses, the effect of game reviews on players is not known. Additionally, the use of game reviews to improve evaluation techniques has received little attention. In this thesis we investigate the effect of game reviews on player experience and perceptions of quality. We show that negative reviews cause a significant effect on how players perceive their in-game experience, and that this effect is a post-play cognitive rationalization of the play experience with the previously-read review text. To address this effect we designed and deployed a new heuristic evaluation technique that specifically uses game reviews to create a fine-grained prioritized list of usability problems based on the frequency, impact, and persistence of each problem. By using our technique we are able to address the most common usability problems identified by game reviews, thus reducing the overall level of negativity found within the review text. Our approach helps to control and eliminate the snowballing effect that can be produced by players reading reviews and subsequently posting their own reviews, and thus improve the commercial success of a game.

ACKNOWLEDGMENTS

The tremendous support and guidance given to me by my supervisor, Regan Mandryk, for the past three years is something that I will be forever grateful for. The work that I have done during my time as a graduate student has only been possible with the careful mentorship that she has given to me. Most importantly, Dr. Mandryk allowed me the freedom to explore a multitude of fascinating research topics. Despite her very busy schedule, Dr. Mandryk always had words of encouragement, and always made time when I needed it. As a consequence, I owe my thanks to Dr. Mandryk for making my experience as a graduate student better than I ever could have hoped for.

I must also thank my wife Sondra and my family for standing by me. Whatever I needed – whether I was tired, unsure, or ready to give up – I could always find the support I needed to keep working.

Finally, I would like to thank everyone whom I have had the opportunity to work with over the last few years in the Interaction Lab. I cannot imagine a better set of friends and colleagues to work with.

DEDICATION

For Sondra, who stands by me patiently and supportively while I chase dreams.

CONTRIBUTIONS OF AUTHORS

The manuscripts found in this thesis were authored in collaboration with four other authors. The manuscripts in chapters 2 and 3 are co-authored by Dr. Regan Mandryk and Dr. Lennart Nacke, and the manuscript in chapter 4 is co-authored by Dr. Regan Mandryk and Dr. Kevin Stanley. The contribution for each author was similar, providing direction in the designing of experiments, sharing domain specific knowledge, and editing and verifying the final manuscripts.

CONTENTS

Permission to Use	i
Disclaimer	ii
Abstract	iii
Acknowledgments	iv
Dedication	v
Contributions of Authors	vi
Contents	vii
List of Tables	xi
List of Figures	xii
CHAPTER 1: Introduction	1
1.1. Game Reviews: More Than Just Buyer Guides	1
1.2. User Experience & Usability	2
1.3. Experience	3
1.4. Cognitive Bias & Mood	4
1.5. Overview	5
1.6. The Manuscripts	8
1.6.1. Manuscripts included in this thesis	8
CHAPTER 2: The Impact of Negative Game Reviews and User Comments on Player Experience	10
2.1. Overview	10
2.2. Introduction	11
2.3. Related Work	12
2.4. Experiment	14
2.4.1. Games	14
2.4.2. Experimental Design and Affective Text Stimuli	15
2.4.3. Participants and Procedure	16
2.4.4. Data Analyses and Dependent Measures	17
2.5. Results	17
2.5.1. Does the valence of text affect the experience?	17
2.5.2. Does the relevance of the affective text to the task matter?	18
2.5.3. Does the authority (expert or peer) of the source matter?	18
2.5.4. Can the valence results be explained by a change in mood as a result of reading positive or negative text?	18
2.5.5. How do our results compare to reading no review text?	19
2.6. Discussion	19
2.6.1. Valence Effects and Review Source	19
2.6.2. Explanations of the Observed Effects	20
2.6.3. Effects on Player Experience	22
2.6.4. Practical and Industrial Considerations	23
2.6.5. Limitations of the Study	24
2.7. Conclusion	24
2.8. Acknowledgments	25
CHAPTER 3: Influencing Experience: The Effects of Reading Game Reviews on Player Experience	26

3.1. Overview	26
3.2. Introduction	27
3.3. Related Work	29
3.4. Experiment	31
3.4.1. Design, Apparatus, Participants, and Procedure	31
3.4.2. Data Processing and Analysis	34
3.4.2.1. Electromyography (EMG)	34
3.4.2.2. Galvanic skin response (GSR)	34
3.4.3. Data Analysis and Dependent Measures	35
3.5. Results	36
3.5.1. Does the text's valence affect perception of game quality?	36
3.5.2. Can the ratings differences be explained by changes in mood?	36
3.5.3. Are the differences reflected in objectively-measured player experience?	37
3.6. Discussion	38
3.6.1. Effects of Review Text on Player Experience	39
3.6.2. Impact on Industry	40
3.6.3. Alternate Explanations of the Results	41
3.6.4. Limitations and Future Work	42
3.7. Conclusion	43
3.8. Acknowledgments	43
CHAPTER 4: Critic-Proofing: How Using Critic Reviews and Game Genres can Refine Heuristic Evaluations	44
4.1. Overview	44
4.2. Introduction	45
4.3. Related Work	47
4.3.1. The Heuristic Evaluation Technique	47
4.3.1.1. Video Game Heuristics	48
4.3.1.2. Severity Ratings	49
4.3.2. Critic-Proofing	50
4.3.3. Video Game Genres	51
4.4. Critic-Proofing Heuristic Technique	53
4.4.1. Technique Overview	53
4.4.2. Pre-evaluation	54
4.4.2.1. Assessing Gameplay & Determining Genres	54
4.4.2.2. Calculate Genre Rating	55
4.4.3. Severity Calculation	57
4.5. Case Study: Capsized	60
4.5.1. The Game	60
4.5.2. A Basic Evaluation	62
4.5.3. Critic-Proofing Heuristic Approach	63
4.5.4. Comparing Results	64
4.5.5. A Developers Reception	66
4.6. Discussion	67
4.7. Conclusion	69
4.8. Future Work	69
4.9. Acknowledgments	70

CHAPTER 5: Conclusion & Future Work	71
5.1. Conclusion.....	71
5.2. Future Work.....	74
5.2.1. Automated System.....	74
5.2.2. Implementation Strategy.....	75
5.2.3. Web Crawler Functionality	75
5.2.4. Data Collection Strategy.....	76
5.2.5. Analysis Strategy	76
5.2.6. Final Thoughts.....	77
References	78
Appendix A.....	83
Affective Text.....	83
A.1. Negative Game Reviews: The Fancy Pants Adventure.....	83
A.1.1. Review 1	83
A.1.2. Review 2	84
A.1.3. Review 3	85
A.1. Negative Game Reviews: Tag: The Power of Paint	86
A.1.4. Review 1	86
A.1.5. Review 2	88
A.1.6. Review 3	89
A.2. Positive Game Reviews: The Fancy Pants Adventure	90
A.1.7. Review 1	90
A.1.8. Review 2	91
A.1.9. Review 3	93
A.3. Positive Game Reviews: Tag: The Power of Paint.....	94
A.2. Review 1	94
A.2.1. Review 2	95
A.2.2. Review 3	96
A.4. Negative Movie Reviews: Vulgar	97
A.2.3. Review 1	97
A.2.4. Review 2	99
A.2.5. Review 3	100
A.5. Negative Movie Reviews: The Singing Forest.....	101
A.2.6. Review 1	101
A.2.7. Review 2	102
A.2.8. Review 3	103
A.6. Positive Movie Reviews: Good Bye, Lenin	105
A.2.9. Review 1	105
A.2.10. Review 2	107
A.2.11. Review 3	107
A.7. Positive Movie Reviews: Amores Perros	109
A.2.12. Review 1	109
A.2.13. Review 2	112
A.8. Negative User Comments: The Fancy Pants Adventure	113
A.9. Negative User Comments: Tag: The Power of Paint	117
A.10. Positive User Comments: The Fancy Pants Adventure.....	121

A.11. Positive User Comments: Tag: Power of Paint	126
Appendix B.....	131
Experimental Material	131
B.1. Demographic Survey Experiments 1 & 2.....	131
B.2. SAM Experiments 1 Survey	133
B.3. Post Play Experiments 1 Survey.....	134
B.4. Post Play Experiments 2 Survey.....	135
B.5. Experiments 2 SAM Survey	135
Appendix C.....	136
Signal Processing MATLAB Scripts.....	136
C.1. GSR Normalization	136
C.2. Baseline Noise Removal.....	136
C.3. EMG Signal Processing.....	136
C.4. EMG Ratio Processing	137

LIST OF TABLES

Table 2.1. Valence Scores for affective text. Columns show the source and valence condition (negative or positive text); rows show the positive or negative LIWC scores for two stimuli.	15
Table 3.1. Means and SD of dependent measures for each phase, split by group. Arousal and valence were measured after the phase on a 9-pt scale, while the physiological measures are averaged over the phase and are normalized between 0-1.	35
Table 4.1 Pinelle's 12 usability problems found in video games.	52
Table 4.2 Normalized Classification Ratings for a slightly modified Nielsen severity ratings. ...	57
Table 4.3 The calculated genre ratings for Capsized based on the Action and Shooter Genres. ...	63

LIST OF FIGURES

Figure 2.1. Means (\pm SE) for user game rating scores by source of affective text and valence of affective text.	17
Figure 2.2. Means (\pm SE) of valence ratings (9-point scale, higher is more pleasant) for pre-text, post-text, and post-game by source and valence (+, -) of affective text.	20
Figure 3.1. Physiological sensors. GSR (left), Zygomatic and Corrugator EMG (right).....	32
Figure 4.1 The 12 common usability problems combined across genre into frequency of occurrence of the problem.	56
Figure 4.2 Gameplay screenshot from Capsized.....	61
Figure 4.3 This graph shows the distribution of usability problems identified in Capsized, October 2009	62
Figure 4.4 A comparison between severity ratings demonstrating the advantage of the prioritized rating.	66
Figure 5.1 System Design.....	75

CHAPTER 1

INTRODUCTION

1.1. Game Reviews: More Than Just Buyer Guides

The reviews of professional critics play an important and influential role in our understanding of all forms of media. While critics have existed for a long time in media such as movies, books, and music, the criticism of games is still relatively young. Game reviews are a mix of technical and artistic domains; both functional and emotional merits must be carefully weighed. As such, game criticism enjoys a high degree of exposure in other media forms such as TV shows, magazines, and websites.

The game reviews themselves contain an abundance of rich information beyond a simple critique of a game. In many ways game reviews play a significant cultural role, not simply evaluating a game on its own but also placing it within the context of all games. Game reviews can be viewed as experience reports from highly experience players, thus acting as rich source of data for both researchers and designers alike. However, game reviews are much more than this. For example, reviews can contain advice for the reader, design suggestions, social discussions, and comparative information. They can comment on business decisions, and inform purchase choices.

The influence that game reviews wield means they affect the commercial success of the games they critique (Joeckel 2007). They do this by having effects beyond simply providing purchase advice to players and consumers. A game that has higher review scores is labeled as a *good game*, and will receive more attention than other games. This attention comes from players, media outlets (websites, TV, news, etc.), and from the game designers themselves. In recent years there has been a move within the game industry as the business as a whole realizes the

effect that game reviews have on the success of a game. As an illustration of this we see developers and publishers making business decisions (such as the amounts of funding to provide for projects or employee bonuses) being directly linked to the critical reception of a game. However, the relationship between the success of a game and reviews might be more complicated than we know. In this thesis we present findings that demonstrate that game reviews do more than inform purchase decisions, they actually affect how a player experiences the game and perceives a game's quality. While it is known that good reviews affect the success of a game, we show that bad reviews actually have a greater effect on the players' experiences with a game.

1.2. User Experience & Usability

Two aspects of play that are commonly considered in game reviews are the user experience and the usability of the game. The importance of user experience and usability has been well studied in many productivity-related domains of interactive systems. We know that usability directly influences productivity, efficiency, and optimal use of productive software. However, video games are by their very nature not productive, efficient, or designed for optimization. Although there are many similarities between what we would describe as *productive software* and video games, there are clearly also vast differences. These differences are the reason that many of our existing methods for evaluating usability and the user experience do not apply directly without some modification. Despite the differences, user experience and usability are of importance to game software because video games are designed to provide players with a positive experience.

The use of usability evaluations is a fundamental and proven strategy to improve the overall user experience of an interactive system. There are a variety of different ways of

evaluating the usability of a system. In this thesis we examine one specific method – heuristic evaluation (Nielsen & Mack 1994). We have established that user experience in video games is fundamentally different from traditional software. However, it is possible to extend what we know of traditional usability methods into the domain of video games. Heuristic evaluation has already been shown to be effective in the evaluation of games (Federoff 2002). As shown by Pinelle, Wong & Stach (2008a), we can extend the traditional heuristic technique to account for the impact of game reviews. A game reviewer can be thought of as an untrained user experience expert with specific domain knowledge, because they spend hundreds of hours playing different games, comparing the relative merits and flaws, and providing critical analysis of why one game works and another does not. They are expert players, much like any user experience specialist is an expert user. Therefore, using game reviews to improve the quality and focus of a heuristic evaluation goes a long way to improve the overall user experience of a video game.

1.3. Experience

The goal of video game play may not be about efficiency or optimization, but video game players do have a concrete goal – to have fun. Video game play tends to produce a more emotional experience than using standard productivity software. As such, when we measure the user's experience in games we limit our effectiveness by using measures such as efficiency, errors, or slips. Instead we must rely on more abstract measures such as fun, engagement, immersion and flow, and emotional measures such as valence and arousal.

We are not the first to connect the measurement of video game user experience and emotional measures. Previous works (Mandryk, et al. 2006) have demonstrated the connection between emotion and player experience. In chapters 2 & 3 we utilize both self-report measures and objective physiological measures of valence and arousal. We measure the influencing effect

that game reviews have on a player's valence and arousal levels. Our goal was to determine whether or not reading emotionally-charged opinions of a game would influence a player's emotional state and perception of a game before, during, and after play. In this thesis we focus primarily on game reviews but in chapter 2 we also examine the effects of user comments. An evaluation of the effects of other external factors, such as marketing or access to sales data, is beyond the scope of this thesis; however we do extend our work with game reviews beyond analyzing their effect on player experience – we also demonstrate that they can be used to improve usability techniques.

1.4. Cognitive Bias & Mood

The term cognitive bias refers to a collection of different effects that all cause a deviation in human judgment during a specific situation. The causes of cognitive biases are diverse but always result in a measurable change in the individual being biased (Kahneman & Tversky 1972). In this thesis we examine four specific cognitive biases: affective priming (Yi 1990), anchoring (Tversky & Kahneman 1974), bandwagon effects (Sundar et al. 2008), and negativity bias (Baumeister et al. 2001). We examine the biasing effect of affective text on the players' perceptions of quality, and their mood measured both in self-report and objectively using physiological measures. While we are careful in our interpretation of our findings – concluding that our observations cannot be attributed to these specific biases – when we combine our results with our arousal and valence measures the effects we observe are likely cognitive in nature, because we observed no change in mood.

Moods in individuals can develop from a prolonged emotional state (Ekman 1994). For example, prolonging the emotional state of happiness will result in the individual developing a happy mood. Previous literature has established that the correlation between the affective

components of arousal and valence and emotional dimensions such as excitement, contentment, distress (Russell 1980). In the experiments presented in this thesis we use the measures of valence and arousal to determine the player's experience. The connection between emotion and experience is demonstrated by Mandryk (Mandryk, Inkpen, et al. 2006). As such, our use of arousal and valence as measures is not unprecedented, and can be considered to be one of the best methods available.

1.5. Overview

This thesis is made up of three manuscripts that have either already been published, or have been submitted for publication. The findings contribute to one of two primary areas. First we present findings from the evaluation of player experience and players' perception of quality. Second we present a usability technique developed for the evaluation of usability in games. Our findings from the evaluation of player experience justify the use of the usability technique we developed – the negative effect of bad reviews on subsequent player experience suggest that game developers should strive to achieve reviews that are positive enough to avoid the previously-described negative snowballing effect. To accomplish this game developers must 'critic-proof' their games. We develop a critic-proofing usability technique specifically designed to discover problems in games that are commonly identified by game reviewers.

Each chapter of this thesis makes a number of unique and important contributions; however, this thesis as a whole makes four primary contributions. First, we demonstrate that exposure to negative game reviews and user comments causes players to rate a game significantly lower than exposure to positive game reviews and user comments, or no reviews or user comments. Second, we demonstrate that this effect is not caused primarily by a change in the player's mood as a result of reading emotionally-charged review text. We measure mood

both with self-report and with objective physiological measures, suggesting that the effect of review text on player ratings of games is primarily cognitive in nature – players decide to like the game less. Third, we present a new heuristic evaluation technique that uses game reviews to create an ordered list of usability problems, which can be used to curb the negative snowballing effect caused by negative game reviews. Forth (and finally) we present a system design that supplements the heuristic evaluation technique and provided valuable validation of the method through data-mining and machine learning. Each chapter also provides supplementary contributions beyond the four primary contributions described above.

In chapter 2 we present five main results. First, we show players who read the negative text rated the game significantly lower than players who read positive text. Second, the reviews or comment text had to relate directly to the game to produce biasing effects. Third, the authority of the source did not create different biasing effects. Fourth, the effects could not be solely attributed to changes in mood as a result of reading the text, as measured using self-report. Fifth, negative text has a larger biasing effect than positive text. These findings are important to this thesis because they demonstrate the effect of game reviews on players. However, these results are also important to the game industry where games are regularly released to professional reviewers prior to official releases.

In chapter 3 we present four important results. First, we confirm our finding from chapter 2 where we observed that reading positive or negative game reviews influenced the players' subsequent ratings of game quality. Second, the difference in perceived quality can't be explained by changes in player valence, but may have been mediated by player arousal. Increases in arousal after playing were significant for players who read positive text and only marginally significant for players who read negative reviews, causing a significant difference between the

two groups of players. Third, we observed significant changes in physiological measures when playing the game compared with reading of the reviews, which illustrates the sensitivity of the measures used. Fourth, there were no differences between players' physiological measures in the different groups of participants. This suggests that the players' experience (as measured objectively) did not change as a result of reading the positive or negative reviews. These findings suggest that the variations in player ratings can be attributed to the cognitive rationalization of the experience with the game reviews, and not a change in the actual experience of the game. The findings also demonstrate the need for a method to evaluate a game prior to release that can account for the negative comments of game reviewers to reduce the biasing effect shown in this and chapter 2. Our results also show us that, while players may experience a game the same way, external factors can influence a player's perception of that experience in a negative fashion.

Chapter 4 provides details on the modified heuristic evaluation technique called critic-proofing. Our technique provides three improvements over the traditional technique. First, we provide a fine-grained prioritized categorization of usability problems. Second, we limit the subjectivity of the assignment of severity ratings by calculating a portion of the rating using a formula based on the characteristics of the game. Third, we show that game reviews - divided by genre - can be used to produce a custom evaluation for any game. The critic-proofing technique demonstrates the effectiveness of using game reviews in an evaluation capacity. We use our critic-proofing technique to evaluate a game that - at the time of publication - was still in development. Feedback collected from the developers suggested that our method was more useful than the traditional heuristic approach. Our results show that despite the effect game reviews can have on a player's perception of quality, we can still use them to correct many critical problems prior to release.

The work in this thesis provides us with both a better understanding of how game reviews affect players, and how they can be used to improve their overall usability. While the main strategy of a reviewer is to inform or persuade a reader of some argument, we show that the reviewer – perhaps unwittingly – provides us with so much more.

1.6. The Manuscripts

As noted previously, this thesis is comprised of three related manuscripts. These manuscripts are outlined below. Each manuscript has been published or accepted for publication at the time of writing this thesis.

1.6.1. Manuscripts included in this thesis

Chapter 2 is built upon the following reference:

(Livingston et al. 2011b)

SIGGRAPH is the premier conference for publishing research related to computer graphics. The SIGGRAPH games track is a top venue to showcase research on games as many of the attendees of SIGGRAPH are in the computer and video games industries. There were 67 papers submitted to the conference track and 8 accepted for an acceptance rate of 12%.

Chapter 3 is built upon the following reference:

(Livingston et al. 2011a).

The international conference on entertainment computing (ICEC) is a top international conference devoted to the academic study of entertainment. Combining research from different domains, such as Computer Science and Game Design, the conference attracts an audience focused on entertainment, but with diverse backgrounds. Held in Vancouver Canada this year, the conference will benefit from the strong game industry present in Vancouver. The acceptance rate for papers submitted to the conference was 31%.

Chapter 4 is built upon the following reference:

(Livingston, Mandryk, et al. 2010)

Future Play is a second-tier conference that brings together researchers from multiple domains that are researching computer games. As one of the only venues to publish academic work on games, it is a growing conference. In 2010 when this paper was published, Future Play was colocated with GDC Canada (Game Developer's Conference) – the top industry conference for game developers – and as such received large attendance and good exposure to the computer games industry.

CHAPTER 2

THE IMPACT OF NEGATIVE GAME REVIEWS AND USER COMMENTS ON PLAYER EXPERIENCE

2.1. Overview

In this chapter we begin with an exploration of game reviews within the context of player experience and perceptions of quality. It is important that we set the groundwork for our understanding of the effects that game reviews and user comments have on players. However, it is also important not to confuse the effect of perceived quality on player experience with the effect on the market success of a game. While the two effects are likely connected given their nature, it would be a mistake to draw this conclusion without further evidence. Thus, in the present chapter, we are careful to present our findings in the context of biased (or influenced) experience. A discussion of the commercial impact is beyond both the scope of this chapter and this thesis as a whole.

The affective text used in the experiment presented in this chapter was analysed using the Linguistic Inquiry and Word Count (LIWC) tool (Pennebaker et al. 2011). LIWC is a text analysis software system that measures the use of different word groups. The word groups used by LIWC are known as dictionaries. Initially, LIWC was developed for the analysis of conversational transcripts. However, more recent studies have used the tool for the analysis of a variety of textual inputs, such as emails, speeches, websites, and blog posts. While LIWC can be used to determine the use of positive and negative emotions, levels of professional and casual language, and personal reference, it can't determine underlining meaning or tone of a text (sarcasm for example). In our experiment we used LIWC as a method to determine the empirical level of positive and negative affect present within our text stimuli.

This chapter presents a study that examines the effects of affective text on the player experience and the perception of game quality. Participants were exposed to a variety of textual stimuli and then required to play a game. Our findings indicate that the effects of reading negative text significantly decreased a player’s perception of the quality of the game played. Together with the findings presented in chapter 3, these results help us understand the effect that game reviews and user comments can have on a player’s experience, which we use to motivate the technique developed in chapter 4. The first step, presented in this chapter, is to understand the effects of reading reviews on players’ perception of game quality.

2.2. Introduction

In games, experience is king. Game reviews can be viewed as expert experience reports giving players an idea of what to expect from the game they want to buy (Larsen 2008). In addition, reviews act as more than just purchase guides, they play a role in shaping our understanding of the reviewed games (Zagal et al. 2009) and of ourselves as gamers (Ivory 2006). Game publishers have acknowledged the importance of reviews for selling their games, and use review aggregation sites to gauge the success of their titles and to determine funding for game franchises. For example, the website Metacritic aggregates reviews from both professional game critics and game players and is used by players to decide whether or not they will purchase a game. However, due to the voice these online review sites provide to game players, the role of professional critics has found competition in consumer word-of-mouth and peer comments. Thus, it is currently unclear who—critics or peers—has the most persuasive power.¹

Due to the prevalence of game reviews and user comments, players form opinions about games before having played them. What remains uncertain is how exposure to these reviews

¹ This chapter contains sections of (Livingston et al. 2011b) which are to appear in the proceedings of Sandbox ’11: The 6th ACM SIGGRAPH.

influences player experience, whether the source (e.g., critic or peer) matters, and whether the player experience results from the player's opinion of the game, or is mediated by a change in their mood. The biasing of attitudes and opinions is not a new topic in social psychology, economics, or marketing. However, biasing caused by textual reviews has received little attention in the fields of game user research and human-computer interaction, especially regarding its effect on player experience.

We present a study examining the effects of review text on player experience. The review text was manipulated to be either positive or negative in tone, either relevant to the subsequent gameplay or not, and derived from either an authority (i.e., critic) or peer (i.e., user). Our results show: 1) Players who read the negative text rated the game significantly lower than players who read positive text; 2) The text had to relate directly to the game to produce biasing effects; 3) The authority of the source (expert or peer) did not produce different biasing effects; 4) The effects could not be solely attributed to changes in mood as a result of reading the text; and 5) Negative text has a larger biasing effect than positive text.

Important for our understanding of biasing effects and player experience in general, our results have particular relevance to the game industry where it is common practice to release games to critics and the media prior to the public release date to get early reviews.

2.3. Related Work

Previous work has shown that negative stimuli have a significantly greater effect on an individual than positive stimuli. In their literature review, (Baumeister et al. 2001) concluded that bad is stronger than good in almost every area of psychology. For example, it has been shown that negative information is more attention grabbing (Baumeister et al. 2001; Fiske 1980), influences impression more, (Fiske 1980), is remembered better (Pratto & John 2005), and is

processed more carefully cognitively than positive information (Baumeister et al. 2001)—resulting in a greater biasing effect of negative stimuli.

The persuasive biasing power of experts and peers has been explored in the domains of movies (Reinstein & Snyder 2005), e-commerce (Sundar et al. 2009), advertising (Yi 1990) and news (Sundar et al. 2007). However, there has only been little investigation in the domain of game research. One exception is a recent industrial study investigating effects of game ratings on perceived game quality and game value (Jenkins et al. 2010). In this study, players were shown a high, low, or no game score and an accompanying comment, such as "Game of the year!" The authors found when players were exposed to a high review score (e.g., 90%) they were more likely to rate a game higher than if they were exposed to a low review score (e.g., 61%). The authors concluded that showing players high game review scores will cause higher player ratings, but attributed their results to anchoring (Tversky & Kahneman 1974). Anchoring² is a biasing effect where an individual's ability to estimate is biased toward a previously known value. Thus, we cannot know whether the results seen in Jenkins et al. (2010) are due only to effects of anchoring, or whether they change the player's perceptions of the game. In our experiment, we eliminated all anchoring effects; we did not present review scores, ratings, or value to players. Instead players only read an affective text stimulus. The use of affective text unrelated to a task to bias participants via a change in their mood is known as affective priming (Yi 1990). We investigate affective priming through the use of movie reviews.

If players reading game reviews are invested in evaluating the information in the review, they will expend cognitive effort in the systematic processing of that information (Chaiken 2003). They may also use cues, such as the length of the review or the review source, to give

² For example, people asked to estimate if a price is more or less than 10 dollars are more likely to guess lower values than those who are asked if a price is more or less than a 1000 dollars.

weight to the information they have read—a mechanic known as heuristic processing (Chaiken 2003; Sundar 2007). Heuristically-processed cues are based on biases players have; for example, that a number of corroborating positive ratings are indicative of quality. The latter is called the bandwagon heuristic: if others think something is good, then I will think that it is good, too (Sundar et al. 2007; Sundar et al. 2009; Sundar 2007). Another related heuristic cue is the authority heuristic: if an expert thinks something is good, I will agree (Sundar et al. 2009).

In the present study, we explore the effect of game reviews, user comments, and movie reviews on player experience.

2.4. Experiment

Our study asks five main research questions:

1. Does the positive or negative tone in review text affect player experience?
2. Does the relevance of the review text to the task matter?
3. Does the authority (expert or peer) of the authors matter?
4. Can the experience differences be explained by changes in mood?
5. How do our results compare to reading no review text?

2.4.1. Games

We used two little-known games: Tag: The Power of Paint, a 3D puzzle shooter, and Fancy Pants Adventures (FPA), a 2D platform game. The primary factor in choosing games was obscurity—the games had to be of professional quality but publicly unknown, so that users did not have prior first or second hand experience with them. No participant expressed prior knowledge of either game when asked after they completed the study.

2.4.2. Experimental Design and Affective Text Stimuli

We used a 3 Source (Game Critic, User Comments, Movie Critic) by 2 Valence (positive, negative) between-subjects design. The sources were either relevant to the task (Game Critic and User Comments) or not (Movie Critic) and were either authoritative (Game or Movie Critic) or peers (User Comments). We also included a control group, who did not read any review text, resulting in seven groups.

We created the review text by modifying real reviews and user comments to ensure that the text stimuli were comparable across conditions and as authentic as possible. For Movie Critic reviews, we modified reviews of obscure films to ensure that participants' response to review text was not influenced by their opinions of the review subject. To create the positive and negative versions, we used affectively-charged antonyms. As an example, the statement "I love this game" in the positive condition would be "I hate this game" in the negative condition. The content and subject matter between conditions was maintained as much as possible. For example, if the positive condition discussed the camera positioning, so did the negative condition.

Table 2.1. Valence Scores for affective text. Columns show the source and valence condition (negative or positive text); rows show the positive or negative LIWC scores for two stimuli.

		Game Critic		Movie Critic		User Comments	
		+	–	+	–	+	–
1	+	7.0	0.8	7.6	2.9	8.5	2.3
	–	1.0	8.0	1.6	6.6	1.0	7.9
2	+	7.6	2.0	8.2	2.3	8.9	2.3
	–	0.5	7.0	1.5	7.6	0.8	7.5

For the Game and Movie Critic reviews, we used multiple corroborating reviews to alleviate participant skepticism surrounding a single review. To increase participant buy-in, they were informed that the text had been taken from a professional review website. In all cases, the

review text was about 1000 words. To ensure that the affective tone (positive or negative) of our text was comparable across conditions, we used the Linguistic Inquiry Word Count (LIWC) text analysis tool (Pennebaker et al. 2011) to measure the valence of review text for each game (see Table 2.1).

2.4.3. Participants and Procedure

Seventy participants (29 female), aged 18–36 ($M=24$) years began by completing a demographic survey, and then were randomly assigned to a group, balancing for sex. Players then completed the experiment using a custom system to deploy the games and gather the survey responses.

Players first provided their pre-text baseline valence rating using the self-assessment manikin (SAM) (Lang 1980), a pictorial 9-point scale commonly used to self-report valence. Players then read the text corresponding to their test group (or read nothing in the control group). They were required to answer questions related to the text to ensure they systematically processed the information, and did not rely solely on heuristic cues (Chaiken 2003). Players then provided their valence rating again, played the game for 15 minutes, and provided a final valence rating. Half of the participants played Tag first, the other half played FPA first. To evaluate their experience, players wrote a short game review and provided a game rating on a 100-point scale. This procedure was repeated for the second game. Before providing their ratings, players were never shown numerical scores to ensure that there were no anchoring effects (Tversky & Kahneman 1974) (when exposure to a numerical rating would influence a player's own numerical rating) as seen in (Jenkins et al. 2010). To help players choose scores, and to aim for consistency across players in their interpretation of the ratings scale, a score guideline was

provided (0–19: I hated it; 20–49: I didn’t like this; 50–74: It was ok, but I’m rather indifferent; 75–89: It was good, but room for improvement; 90–100: I loved it).

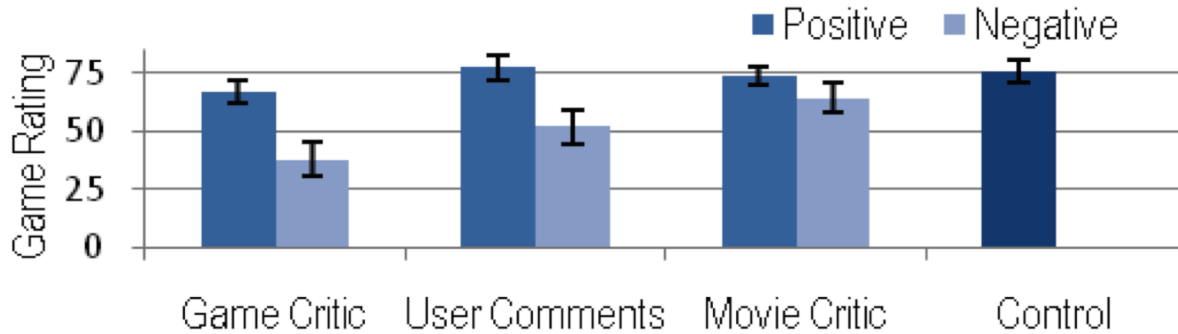


Figure 2.1. Means (\pm SE) for user game rating scores by source of affective text and valence of affective text.

2.4.4. Data Analyses and Dependent Measures

Our dependent measures are user ratings on a scale of 1–100 and user valence ratings from the SAM. These do not meet the assumptions of normality and equal variance measured using the Shapiro-Wilk test for normality. As such, we used the Mann-Whitney Test for all between-groups tests, and the Wilcoxon Signed Ranks Test for within-player comparisons (Huck & Cormier 1996). Only data from the first game played are presented, because initial analysis showed that players consistently discovered the experimental manipulation after playing the first game and answering the surveys, biasing the data from the second game. Consistent with previous work (Chaiken 2003; Jenkins et al. 2010; Chaiken & Maheswaran 1994), we discarded the data from the biased game trial.

2.5. Results

2.5.1. Does the valence of text affect the experience?

Players who read positive text rated the game higher than players who read negative text ($Z = 3.6, p \approx .000, M^+ = 72.9, SE^+ = 2.6$) ($M^- = 51.3, SD^- = 4.4$).

2.5.2. Does the relevance of the affective text to the task matter?

Tests for each text source (game critic, user comments, and movie critic) showed that reading positive text resulted in higher ratings than reading negative text for the players reading game critic reviews ($Z = 2.8, p = .004$) and user comments ($Z = 2.5, p = .011$) for the game that they subsequently played, but not for players reading movie critic reviews ($Z = 1.0, p = .315$). See Figure 2.1.

2.5.3. Does the authority (expert or peer) of the source matter?

The experience ratings for players reading negative game critic (expert) and user (peer) text were not significantly different ($Z = -1.5, p = .127$). Neither were ratings for players reading positive game critic and user text ($Z = -1.7, p = .081$). See Figure 2.1.

2.5.4. Can the valence results be explained by a change in mood as a result of reading positive or negative text?

We were interested in determining whether reading the affective text had a direct impact on player experience, or whether it caused a change in players' overall affective state (mood), which then had an impact on player experience. We had asked participants to rate their mood both prior to and directly after reading the affective review text. We compared these two ratings for each of the six groups (3 Source \times 2 Valence) and found that none of the differences were significant except for those who read positive movie text ($Z = 2.1, p = .034$). As Figure 2.2 shows, reading the positive movie review text improved players' self-rated valence. All other differences failed to reach significance ($p_{game-} = .666, p_{user-} = .103, p_{movie-} = .336, p_{game+} = .085, p_{user+} = .230$).

2.5.5. How do our results compare to reading no review text?

We included a control group that read no review text before playing the game to see whether the positive reviews were improving play experience or the negative reviews were worsening play experience. Players reading negative affective text rated the game lower than players reading no text at all ($Z = 3.0, p = .002$), but players reading positive affective text rated the game no differently than players reading no text at all ($Z = 1.2, p = .259$). See Figure 2.1.

2.6. Discussion

There are several results to take away from this study:

1. Reading negative review text negatively influences player experience.
2. To affect player experience, the text must relate to the game (e.g., a review of the game played).
3. The authority of the source does not matter; the peer and expert sources did not differ significantly in their influence on experience.
4. There was no significant difference in mood for players before and after reading game reviews and user comments.
5. The positive review text provides no positive benefit over no review text at all.

We next discuss the meaning of our findings and how they compare to prior work, their industrial relevance, and future research directions.

2.6.1. Valence Effects and Review Source

In our experiment we showed that the valence of review text significantly affected player experience. We further show that the negative text had a greater effect on player experience, and that the authority of the related review source (critic versus peer) does not matter. Interestingly, our results suggest that the observed effect on player experience was not primarily mediated by changes in player valence. If this were the case, reading the affective text would produce a

change in mood, which would then produce a resulting change in player experience. Our results show that although the differences in player ratings were significant between the groups who read positive or negative reviews, the reviews did not produce a consistent and corresponding affective state change. It could be that the SAM ratings data was too noisy to find significant differences, but Figure 2.2 shows no consistent pattern, thus we are quite confident that the ratings differences cannot be attributed to changes in overall mood.

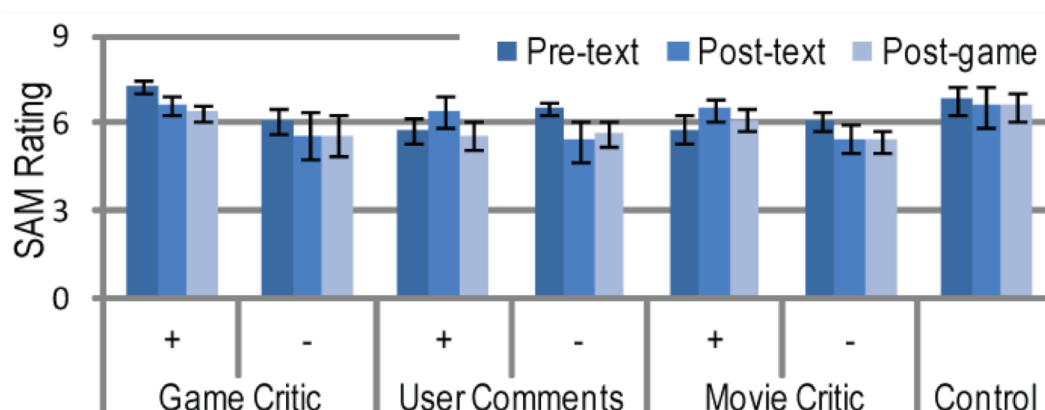


Figure 2.2. Means (\pm SE) of valence ratings (9-point scale, higher is more pleasant) for pre-text, post-text, and post-game by source and valence (+, -) of affective text.

Our findings also suggest a primary contributing factor is relevance of the text to the game being played. The game reviews and user comments about the game produced a consistent and significant difference in player game ratings, while the movie reviews did not.

2.6.2. Explanations of the Observed Effects

We believe that systematic processing of the review text influenced players' opinions of and experience with the game; however, previous literature has suggested that changes in experience could be explained by heuristic cues, anchoring, or affective priming. We address each of these in turn.

The *bandwagon heuristic* suggests that if others think that something is good, then I will agree, while the authority heuristic suggests that if an expert thinks something is good, then I will agree. Previous literature on the effects of experts and peers (Sundar 2007; Sundar et al. 2009) has shown the authority heuristic can be superseded by a bandwagon heuristic. Our results showed no differences in player experience after reading expert reviews or user comments. It is possible that systematic processing of the negative review text caused the prioritization and interpretations of negative content to overshadow any authority or bandwagon heuristic effects. In addition, we did not observe a difference between the control condition and any of the positive conditions, suggesting that when information is systematically processed, heuristic cues are less effective.

Affective priming uses a stimulus, unrelated to the task or object to be viewed, to change a participant's affective state or mood (Yi 1990). In our movie review condition, we did see a positive impact of a positive review on players' moods, but not a negative impact of a negative review. However, there was no effect on player experience, thus our results cannot be explained by affective priming. Because the review text had to be related to the game to show effects, we can assume that task-relevance has a much stronger effect on player experience than affective priming.

Anchoring is a biasing effect where an individual's ability to estimate is biased toward a previously-known value. In previous work, players exposed to a negative rating of a game rated that game lower themselves (Jenkins et al. 2010). This is clearly an effect of anchoring, so the previous work does not show that player experience was actually affected. In our study, we did not present the players with any numerical ratings, so the content and tone of the review text was causing the difference in player experience. Thus, our results cannot be attributed to anchoring.

We conclude it is cognitive processing and interpretation of negative text that has the greatest effect on player experience, not anchoring, affective priming, or heuristic cues. However, additional work examining the methods and effects of information processing, is required to confirm this.

2.6.3. Effects on Player Experience

Although we observed changes in game ratings as a result of reading affective reviews, the question remains whether the ratings differences were a result of a change in actual player experience or a reflection of the player's cognitive rationalization of their experience with the review text. In the former situation, a player who read a negative game review would actually not enjoy the experience as much as a player who read a positive review. In the latter, both players would have a similar experience; however, the one who read a negative review would interpret their experience as having been more negative.

To consider this, we can look at how players' moods changed over the course of gameplay. A significant correlation between review score and post-play SAM ratings (Spearman's $\rho = .442$, $p \approx .000$) show that the player's self-rated affective state and their rating of the game correspond. So although reading the review text failed to produce an observable change in players' affective state, playing the game after reading the review text did. The more positive that participants were feeling after playing, the higher they rated the game. This result leads us to think that actual experience was changed, but to better explore this question we must measure player experience during play. We plan to investigate this issue by measuring player experience at the physiological level (Mandryk 2008), which would tell us if the biasing effects from affective reviews are mirrored physiologically during play or if the mind tells a different story than the body.

2.6.4. Practical and Industrial Considerations

Compared to no affective text, negative reviews and user comments had a detrimental effect on player experience, whereas positive reviews produced little to no relative improvement. These results have practical significance for game designers, online community managers, and marketers.

Unlike other products, games provide entertainment experiences that not only require an investment of money but also of personal leisure time, which makes an informed purchase decision important for players. Game reviews facilitate this decision by providing player experience reports, giving players an idea of what to expect from the game they want to buy. It is standard practice for developers and publishers to show a game to media outlets prior to its release. If the game is well received, it can generate hype, which can translate into sales. Our results suggest that reading a positive review does not provide benefit in terms of player experience over reading no review, but that a negative review can significantly harm player experience, which would likely generate a negative hype effect. Therefore, negative reviews should be avoided, for example through a critic-proofing approach to playability testing (Livingston, Mandryk, et al. 2010).

In an online community this negative effect can snowball as players who are influenced by critic reviews may in turn post negative comments online, influencing an even larger player population. As this cycle continues with new players reading and being influenced by the negative press, the bottom line of a game in terms of sales and future sequels could be negatively affected.

2.6.5. Limitations of the Study

Our study measured player experience using game ratings and their affective state using the SAM, so we can only report on players' perceived experience. It is possible that all players had a similar game experience, but that those exposed to negative reviews perceived their experience as more negative due to the influence of the affective text. The question is whether player experiences were different, or whether their interpretations of the experiences were different. We plan to investigate this issue by measuring player experience at the physiological level (Mandryk 2008), which would tell us if the biasing effects from affective reviews are mirrored physiologically during play or if the mind tells a different story than the body.

2.7. Conclusion

Game reviews and player ratings have an effect on the commercial success of games, but it has not been clear how game reviews affect player experience. We investigated how positive and negative text influences players' affective states and their play experiences. Our results show: 1) Players who read the negative text rated the game significantly lower than players who read positive text; 2) The text must be relevant and related to the game to produce significant player experience effects; 3) The authority of the text (expert or peer) did not influence the biasing effect; 4) The effects could not be explained by changes in mood as a result of reading affective text; and 5) Negative affective text has a larger biasing effect than positive affective text, likely due to the more careful processing of the negative information. Important for our understanding and analysis of player experience, our results have particular relevance to the game industry where it is common practice to release games to critics and the media prior to the public release date.

2.8. Acknowledgments

This research was funded by NSERC and the GRAND NCE. Thanks to David McDine for literature suggestions and our past reviewers who provided helpful comments on previous versions of this document. We are also grateful to all participants for taking part in this study.

CHAPTER 3

INFLUENCING EXPERIENCE: THE EFFECTS OF READING GAME REVIEWS ON PLAYER EXPERIENCE

3.1. Overview

In chapter 2 we demonstrated the impact that negative game reviews and user comments can have on perceived quality of a game; however, there still remain a number of unanswered questions. Specifically the study presented in chapter 2 does not evaluate the player experience directly, so it is unknown if the observed effects on game ratings were primarily cognitive or whether the reading of affective text actually causes an observable change in the player's experience of the game.

We address this question in this chapter by extending our work and findings from chapter 2. In the current chapter we present a study where we measure a player's mood using both self-report and objective physiological measures. To determine the effects of reviews on player experience and players' perceptions of game quality, we designed an experiment that measures the physiological response of a player both during the reading of the affective text and during the playing of the game. The correlation of player experience with galvanic skin response (GSR) and electromyography (EMG) has previously been established by (Mandryk, Inkpen, et al. 2006). We know (because of the relationship between these physiological measures and valence and arousal) that a change in mood suggests a change in experience. The analysis of our results yielded a number of interesting findings supporting our hypothesis that the effects of affective text are primarily cognitive.

While our findings in this chapter are important for our general understanding of the effects of game reviews, they also demonstrate the value of game reviews for evaluation purposes. In chapter 2 we suggested that the reading of negative reviews or user comments might

cause a snowballing effect, where the negativity cascades as more individuals read the reviews. It should be noted that we did not observe this effect directly; however, based on prior literature the effect is certainly likely. The results from chapter 2 & 3 suggest that game developers should strive to produce games that are *good enough* to garner positive reviews so as to eliminate – or at least limit – the snowballing effect created by negative game reviews or user comments.

3.2. Introduction

Game reviews play a major role in the video game industry and in video game culture. They shape our understanding of games, sway business decisions and consumer choice, and impact player experience. The combination of these and other factors has made video game reviews a large and influential business. Players form opinions about a game before ever playing them due, in part, to the influence game reviews wield over other aspects of the business such as marketing, and also directly over players' purchase intent.¹

In our previous work (Livingston et al. 2011b) we demonstrated that biasing effects from reading negative game reviews cause players to rate the quality of a game significantly lower than after reading positive reviews, or reading no reviews at all. Our results showed that the changes in perceived game quality were not mediated by subjective changes in the players' overall moods. However, what remained unclear was whether the effect of game reviews caused a change in the player experience, or if we observed a post-play cognitive rationalization of player experience with the content of the review text. In the former situation, a player who read a negative review would enjoy the play experience less than a player who read a positive review. In the latter, both players would have a similar experience; however, the one who read a negative review would interpret their experience as having been more negative. This distinction is

¹ This chapter contains sections of (Livingston et al. 2011a), which are to appear in the proceedings of ICEC 2011

important as it helps us understand the impact of reading game reviews on subsequent player experience in a time when it is common for players to read about a game prior to making a purchasing decision or before even playing a demo version of the game.

To determine whether player experience is directly affected by review text or whether the biasing effects are a cognitive rationalization of player experience with the content of the review, we designed an experiment that measures a player's subjective rating of a game, their subjective mood throughout the experiment, and their objective play experience throughout the experiment. To provide an objective measure of play experience, we collected a player's physiological data during game play. It has previously been shown that the body's physiological reaction changes when presented with a stimuli – such as a video game – and correlates strongly with the user's emotional experience (Mandryk et al. 2006) and in-game events (Hazlett 2008). We operationalize objective player experience through the measurement of two orthogonal signals – galvanic skin response (GSR) to measure arousal (degree of activation), and facial electromyography (EMG) to measure valence (positive or negative). Differences in the physiological measures would reveal that players reading negative reviews experience the game differently from players reading positive reviews, while similar measures would reveal that the play experiences themselves are not different.

Similar to our previous research, our results show that players' ratings of a game are systematically affected by reading positive or negative review text of a game prior to playing it. We also show that these differences are not fully explained by subjective changes in the players' mood as a result of reading emotionally-charged text. As predicted by previous work, we found significant physiological differences for the different tasks in the experiment (i.e., reading reviews, playing game). Finally, we did not find objective differences in players' physiological

response to the game depending on the valence of the review text for either the valence or the arousal signals. These findings suggest that the variations in player ratings of the game's quality can be attributed to a cognitive rationalization of their experience with previously read review text, and not to differences in the play experience itself. Important for our understanding of the impact of reading game reviews on subsequent player experience, we have shown that despite rating the games differently, players' experience of the game was essentially the same.

3.3. Related Work

Previous work has shown that negative stimuli have a significantly greater effect on an individual than positive stimuli, an effect that has been observed in almost every area of psychology (Baumeister et al. 2001). For example, it has been shown that negative information is more attention grabbing (Fiske 1980; Baumeister et al. 2001), influences impression more (Fiske 1980), is remembered better (Pratto & John 2005), and is cognitively processed more carefully than positive information (Baumeister et al. 2001), resulting in a greater biasing effect of negative stimuli.

Persuasive power has been explored in the domains of movies (Reinstein & Snyder 2005), e-commerce (Sundar et al. 2009), advertising (Yi 1990), and news (Sundar et al. 2007); however, there has been little investigation into the effect on player experience. Jenkins et al. (2010) studied the effects of showing a high, low, or no game score to players. The authors found when players were exposed to a high review score (e.g., 90%) they were more likely to rate a game higher than if they were exposed to a low review score (e.g., 61%), and attributed their results to "anchoring" (Tversky & Kahneman 1974). *Anchoring* is a biasing effect where an individual's ability to estimate is biased toward a previously known value. In our previous work (Livingston et al. 2011b) we found that the biasing effect remained when anchoring effects were

removed. We presented participants with various affective text stimuli including game reviews, user comments, movie reviews, and no text at all. We found that the valence of the reviews significantly affected player ratings, that the authority of the text did not matter (both user comments and critic reviews showed effects), and that the relevance of the text did matter (no differences were observed after reading positive and negative movie reviews). We also found that the biasing effect was strongest in the negative text conditions – consistent with similar work in psychology (Baumeister et al. 2001) – and likely primarily a cognitive effect caused by the more careful processing of the negative affective text. Finally, we showed that the ratings differences could not be explained by changes in subjectively-measured mood. This work raised the question of whether the ratings differences were a result of a change in player experience or a reflection of the player’s cognitive rationalization of their experience with the review text. The present study answers this question by evaluating the effect of game reviews on objective player experience measured using physiological signals.

The study of players’ physiological responses to video game stimuli has seen a significant amount of research. Mandryk et al.’s work (Mandryk, Atkins, et al. 2006; Mandryk, Inkpen, et al. 2006) has explored the psychophysiological effects of games on an individual, and showed that the physiological measures were related to the subjective responses of players (Mandryk, Inkpen, et al. 2006), and that the physiological measures could be used to build a mathematical model of player emotional state (Mandryk, Atkins, et al. 2006). Drachen et al. (2010) also found a significant correlation between psychophysiological arousal and self-reported player experience. Hazlett showed a relationship between players’ physiological measures and in-game events. He found positive facial EMG activity to be higher during positive events, while negative facial EMG activity was greater during negative events (Hazlett 2006;

Hazlett 2008). Ravaja et al. (2006) also showed a relationship between in-game events and phasic responses in GSR, EMG and heart rate acceleration, and Nacke et al. (2010) showed that challenging game situations evoke positive affect and high arousal in players as measured by GSR and EMG. Finally, although there are significant individual differences in physiological responses, researchers have demonstrated success in the normalization of GSR (Mandryk 2008) and EMG (Hazlett 2008) data for between-participant comparisons. These works demonstrate the relationship between physiological responses with both self-report and in-game events. However, these studies focus on the effects of the game condition or the game itself on physiological measures. In our study, we extend these works by examining the effect of systematically-processed (Chaiken & Eagly 1989) information read prior to play on a player's physiological response.

3.4. Experiment

3.4.1. Design, Apparatus, Participants, and Procedure

We used a single-factor between-subjects design manipulating two levels of valence (positive, negative). In our previous study we found the biasing effect manifests when the text is relevant to the task being performed (Livingston et al. 2011b), so we used game review text. We created the review text by modifying real reviews to ensure that the text stimuli were comparable across conditions and as authentic as possible. To create the positive and negative versions, we used affectively-charged antonyms. To control heuristic cue effects (Chaiken & Eagly 1989) the content and subject matter between conditions was maintained as much as possible. To increase participant buy-in, they were informed that the text had been taken from a professional review website. In all cases, the review text was about 1000 words. To ensure that the affective tone (positive or negative) of our text was comparable across conditions, we used the Linguistic

Inquiry Word Count (LIWC) text analysis tool (Pennebaker et al. 2011) to measure the valence of review text for each game (see Livingston et al. 2011b).

For our experiment we used a little-known game. Tag: The Power of Paint is a 3D puzzle shooter, and winner of IGF's 2009 Best Student Game. The primary factor in choosing a game was obscurity – the game had to be of professional quality, but be unknown so that users did not have prior first or second hand experience with it. All participants were unfamiliar with Tag prior to participating in the experiment.

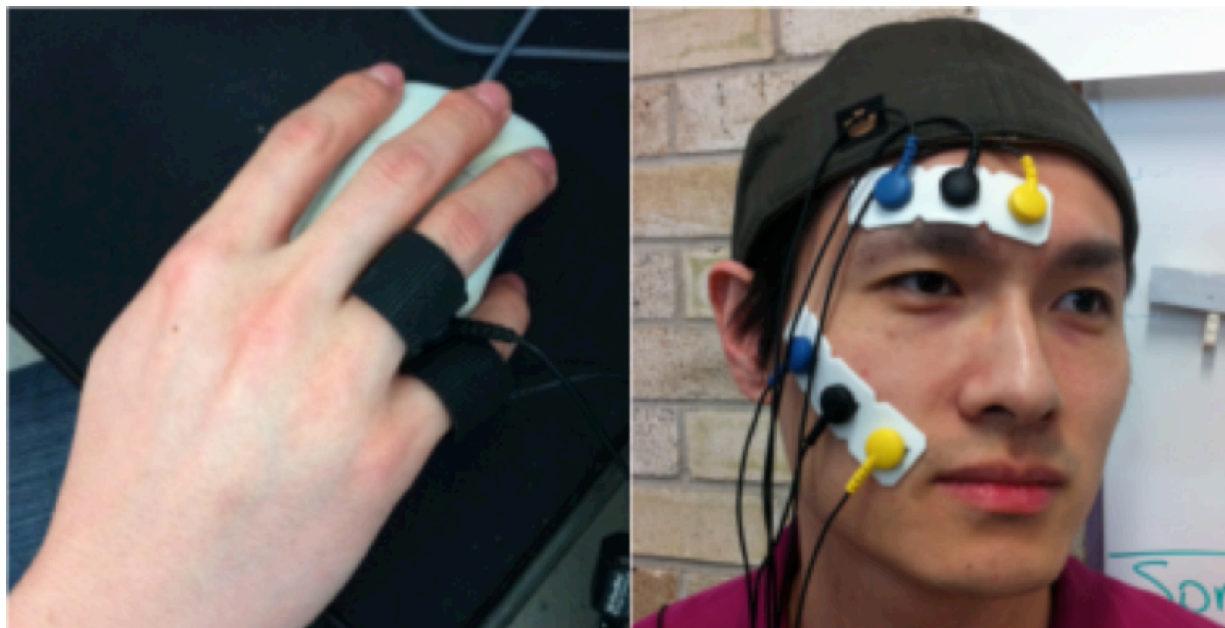


Figure 3.1. Physiological sensors. GSR (left), Zygomatic and Corrugator EMG (right).

To measure physiological responses, we used Thought Technology sensors (EMG, GSR) connected to the FlexComp Infiniti encoder. The EMG band sensors were attached to the right corrugator and zygomatic muscles of the face (Figure 3.1). To ensure strong EMG signals, players were screened prior to participation for excessive facial hair. The GSR sensors were attached to the ring and pinky fingers of the right hand (Figure 3.1). The right hand was chosen because players moved the fingers less using the mouse with the right hand than using the keyboard with the left.

Twenty participants (8 female), aged 18-25 ($M=21.8$ years) began by completing a consent form, and were then randomly assigned to one of the two experimental groups, balancing for sex. Players completed the experiment using a custom system developed to deploy the game and gather the survey responses using an online form.

Players were first connected to the physiological sensors and the setup was tested (for about 1 minute) to ensure strong signals from each sensor. Players then provided their demographic information. Physiological data was collected at four times during the experiment: during the baseline, during the reading of textual stimulus, while answering stimulus questions, and during gameplay. The baseline and gameplay sessions were timed to record precisely four and fifteen minutes respectively. Players were not time-limited during the reading or answering phases, rather markers were added to the data when the player started reading, stopped reading, started answering, and stopped answering. During the baseline phase, players were shown a gray screen with a cross displayed slightly off center. The player then performed a pre-text valence and arousal rating via the self-assessment mannequin (SAM) (Lang 1980). Next, participants read the text stimulus corresponding to their test condition and answered the required questions that followed. The SAM valence and arousal was measured again, and then the game was played for 15 minutes. Afterwards, the SAM valence and arousal ratings were collected a final time.

To evaluate their perception of game quality, players wrote a short game review and provided a game rating. Players were never shown numerical scores to ensure that no biasing anchoring effects were present. To help players choose scores, and to aim for consistency across players in their interpretation of the ratings scale, a guideline was provided (0-19: I hated it; 20-49: I didn't like this; 50-74: It was ok, but I'm rather indifferent; 75-89: It was good, but room for improvement; 90-100: I loved it).

3.4.2. Data Processing and Analysis

Physiological data were exported from BioGraph as raw CSV files and imported into MATLAB for processing and analysis. Data were divided based on the action being performed. Subsequent analyses were conducted during reading of the stimulus text and playing the game, although the baseline phase was used for signal normalization. EMG and GSR require different signal processing, so custom scripts were developed for each physiological sensor.

3.4.2.1. Electromyography (EMG).

The raw data collected from both EMG sensors was the RMS of the signal. Consistent with normal EMG signal analysis (Tassinari & Cacioppo 2000) the raw data was first filtered, and then smoothed. All EMG data was subjected to a 3rd order Butterworth low-pass filter at 500Hz, then smoothed using a moving average window of 0.5s (1024Hz). The remaining data points were normalized by subtracting the baseline or reading values (whichever were lower) from each sample and representing each sample as a percentage using the min and max values found across all filtered and smoothed samples for a specific participant. Normalizing was required to allow for meaningful between-participant comparisons.

An EMG ratio was also calculated using a similar approach to the one described by Hazlett (2008). We calculated the total time spend in activation higher than one standard deviation above the mean for both the corrugator and zygomatic muscles. The ratio of zygomatic to corrugator activation is expressed as a decimal value. These ratios can safely be compared between participants.

3.4.2.2. Galvanic skin response (GSR).

Similar to the EMG data processing, the GSR values need to be normalized to allow for meaningful statistical comparisons between participants. GSR was normalized by calculating

each sample as a percentage of the entire span of GSR, using the min and max values over all samples for one participant. This practice is consistent with GSR analysis of gameplay experience (Mandryk 2008). The mean and standard deviation for each GSR sample was calculated and recorded as a percentage.

3.4.3. Data Analysis and Dependent Measures.

Our subjective dependent measures are SAM arousal, SAM valence, and game ratings. Our objective dependent measures are normalized GSR, normalized zygomatic and corrugator EMG, and Hazlett EMG emotion ratio. The subjective dependent measures do not meet the assumptions of normality and equal variance, which is typical with ratings data. We used the Mann-Whitney Test for all between-groups tests, and the Wilcoxon Signed Ranks Test for within-player comparisons (Huck & Cormier 1996). The objective physiological measures were analyzed using the appropriate parametric statistical tests.

Table 3.1. Means and SD of dependent measures for each phase, split by group. Arousal and valence were measured after the phase on a 9-pt scale, while the physiological measures are averaged over the phase and are normalized between 0-1.

	Positive Review			Negative Review		
	Baseline	Reading	Playing	Baseline	Reading	Playing
Ratings			80.9(8.61)			44.3(22.94)
Arousal	2.9(0.88)	3.3(1.77)	6.1(1.66)	2.6(0.70)	2.7(1.06)	3.9(1.85)
Valence	6.2(1.55)	6.2(1.62)	7.1(1.45)	7.0(0.94)	5.8(1.32)	6.1(2.33)
EMG Cheek		0.05(0.06)	0.09 (0.09)		0.02(0.04)	0.11(0.07)
EMG Brow		0.04(0.03)	0.01 (0.03)		0.02(0.03)	0.02 (0.02)
GSR		0.23(0.13)	0.65(0.08)		0.23(0.20)	0.67(0.07)
EMG Ratio		1.36(0.86)	1.67(0.93)		1.03(0.53)	2.10(1.47)

3.5. Results

3.5.1. Does the text's valence affect perception of game quality?

Consistent with our findings from our previous study, players who read positive text rated the game significantly higher than players who read negative review text ($Z = 3.6$ $p \approx .000$, $M^+ = 80.9$, $SE^+ = 2.7$, $M^- = 44.3$, $SE^- = 7.3$).

3.5.2. Can the ratings differences be explained by changes in mood?

Consistent with our previous work, we were interested in whether the game ratings differences could be explained by mediating changes in a player's mood. Perhaps the players reading positive text were put into a more positive mood and then enjoyed the game more than players reading negative text (thus rating it higher). We asked players to rate their valence and arousal three times: prior to reading the game review, after reading the game review, and after playing the game. Means and variances are in Table 3.1.

There were no differences between the two groups for any of the three valence ratings ($Z(p)$ - Pre-read: 1.07(.315); Post-read: .70(.529); Post-play: .97(.353)). There were also no differences in arousal between the groups prior to or after reading the game reviews ($Z(p)$ - Pre-read: .80(.481); Post-read: .82(.481)). However, there was a difference in the arousal rating after playing the game, where the positive review group rated their post-game arousal significantly higher than the negative review group ($Z(p)$ - Post-play: 12.45(.015)). To determine what this difference means, we looked into arousal changes over the experiment for each group. There were two arousal differences that we considered – the difference between before and after reading the review text and the difference between after reading review text and after play. For the negative group, there was no difference in arousal after reading the review text ($Z=.58$, $p = .564$) or after play ($Z = 1.78$, $p = .076$). For the positive group, arousal did not increase after

reading the review ($Z = .96, p = .336$), but did increase after playing the game ($Z = 2.68, p = .007$). Thus it appears that reading the review text did not increase players' arousal for either group; however, playing the game significantly increased the positive group's arousal, but only marginally increased the negative group's arousal, resulting in a net difference between the groups.

3.5.3. Are the ratings differences reflected in objectively-measured player experience?

The main goal of the present study was to determine whether the game ratings differences are reflected in differences in player experience during play, or whether the differences are a cognitive rationalization of similar play experiences with differing review texts. We conducted a RM-MANOVA with one between-subjects factor (valence of review – positive or negative) and one within-subjects factor (phase of experiment – reading review or playing game) on four dependent measures (GSR, EMG cheek, EMG brow, EMG ratio). Means and variances are shown in Table 3.1.

As expected, there was a significant main effect of phase on GSR ($F_{1,18} = 143.6, p \approx .000, \eta^2 = .89$), EMG cheek ($F_{1,18} = 15.0, p = .001, \eta^2 = .46$), and EMG ratio ($F_{1,18} = 6.2, p = .023, \eta^2 = .26$), which showed that players were more aroused and experienced higher valence during game play than during reading. There was no effect of phase on EMG brow ($F_{1,18} = 1.0, p = .329, \eta^2 = .05$). These results confirm our expectations that gameplay is more exciting than reading review text, and also demonstrate the capability of the physiological signals to differentiate player experience objectively.

There was no main effect of valence of review text on any of the dependent measures (GSR: $F_{1,18} = .03, p = .879, \eta^2 \approx .00$; EMG cheek: $F_{1,18} = .08, p = .780, \eta^2 \approx .00$; EMG brow: $F_{1,18}$

= .75, $p = .399$, $\eta^2 = .04$; EMG ratio: $F_{1,18} = .02$, $p = .885$, $\eta^2 \approx .00$). These results show that there were no overall objective differences between the two groups.

There were also no interaction effects between valence group and phase on any of the dependent measures (GSR: $F_{1,18} = .021$, $p = .886$, $\eta^2 \approx .00$; EMG cheek: $F_{1,18} = 2.0$, $p = .175$, $\eta^2 = .10$; EMG brow: $F_{1,18} = 1.9$, $p = .184$, $\eta^2 = .10$; EMG ratio: $F_{1,18} = 1.9$, $p = .187$, $\eta^2 = .10$). These results show that the two groups did not respond differently during the two phases. We may have expected similar levels of objective experience during reading, and that the positive group might demonstrate more enjoyment during play than the negative group. However, this is not the case as our results suggest that there are no differences in the response of the two groups to the experiment.

3.6. Discussion

There are four main results to take away from our findings:

1. Reading positive or negative game reviews influences subsequent player ratings of game quality.
2. The ratings differences cannot be explained by changes in the self-reported player valence, but may have been mediated by player arousal. Increases in self-reported arousal after playing were significant for players who read positive reviews, and marginal for players who read negative reviews, resulting in a between-group difference in self-reported arousal post play.
3. There were significant increases in GSR, EMG of the cheek, and the Hazlett EMG ratio when playing the game over reading the review text, demonstrating the sensitivity of these signals to measure player experience.

4. There were no differences in any of the physiological measures between those who read positive or negative reviews, suggesting that player experience did not change as a result of reading review text, but that ratings differences were a post-play rationalization of experience with the previously-read reviews.

The first three results are consistent with previous work, while the fourth provides resolves a question left unanswered by previous work. We focus our discussion on the contributions of game-related pre-play and play experiences on resulting opinion about a game, the potential impact of our results on industry practices, and the research opportunities in this space.

3.6.1. Effects of Review Text on Player Experience

We conducted this work to shed light on an unanswered question – does reading positive or negative game reviews affect player experience? Previous work had shown that reading positive or negative game reviews affected player perception of game quality provided by a game rating, and showed that these differences were not mediated by changes in a player’s mood (Livingston, et al. 2011b). Our work replicated these results, but also showed that reading positive or negative game reviews did not affect player experience during play of the game itself. So if players who read positive text did not experience the game differently than those who read negative text, why did they rate the game significantly higher?

We feel that the game ratings are not indicative solely of the player’s experience, but are actually a cognitive rationalization of the experience in the context of the previously-read review text. This does not mean that players are under a false impression of their experience, or that review text wins when there is a discrepancy between the review content and their experience. It is possible, and quite likely, that players who read negative text would subjectively characterize

their experience as worse than players who read positive text, indicating that the rationalization of experience with the influencing reviews affected not only the perception of game quality, but also the perception of the play experience. That measured player experience did not differ does not make the reflective difference in experience any less real to a player or any less true. It does mean that players who read negative reviews decided that they liked the game less than those who read positive reviews, not because of differences in play experience, but in spite of no differences in play experience. This lack of difference in play experience after reading game reviews has important impact for the game industry.

3.6.2. Impact on Industry

Current review methods used in the game industry include game reviews from both professional and amateur critics, and discussion forums where players can discuss their opinions and experiences. These textual channels offer players a freely available source of influential reports, which influence player perception of game quality. The game industry is known for releasing games to critics prior to release. As our results show that reading reviews (especially negative reviews) biases the player's perception of the game, game studios should be careful of critical receptions prior to game release. It is tempting to believe that a good play experience will outweigh negative buzz surrounding a game, but we show that this may not be the case. Even though experience of the game was not different, players who read negative reviews were still more critical in their perception of the game's quality.

Our findings can also be extended to other marketplaces where the exposure of reviews and ratings is becoming more prominent. For example, the Apple app store prominently displays user ratings and comments right on the main app page for each app. Our results suggest that these ratings may adversely affect the perception of quality of those products for users who read

the reviews prior to using the app. Although these ratings often come from users of the apps and not professional critics, our previous work revealed that the authority of the review source did not matter; players were influenced both by game reviews and by user comments. Developers need to be aware of the impact these reviews and comments have, not only directly on their sales, but also on the experience users have with their products.

In an online community this negative effect can snowball as users who are influenced by critic reviews may in turn post negative comments online, influencing an even larger user population. As this cycle continues with new users reading and being influenced by the negative press, the bottom line of a game or an app in terms of commercial success could be negatively affected.

3.6.3. Alternate Explanations of the Results

There are cognitive biases that can influence a user's perception of a game quality. We controlled for related biases in our study in order to attribute the results to reading of review text and not other cognitive biases. For example, affective priming uses a stimulus, unrelated to the task or object to be viewed, to change a participant's affective state or mood (Yi 1990). Because our review text was related to the game, we can assume that our results are not due to affective priming. Anchoring is a biasing effect where an individual's ability to estimate is biased toward a previously-known value. In previous work, players exposed to a negative rating of a game rated that game lower themselves (Jenkins et al. 2010). The authors identify this as an effect of anchoring. In our study, we did not present the players with any numerical ratings, so the content and tone of the review text was causing the difference in perception of game quality. Thus, our results cannot be attributed to the biasing effects of anchoring or affective priming.

Although players rated the game differently depending on the valence of previously-read review text, we observed no physiological differences between the groups. One might suggest that there were in fact differences in player experience, but that physiological measures are not sensitive enough to show differences. Because our results showed significant and predictable physiological differences between reading the review text and playing the game, we feel that the signals are a good measure of player experience and are sensitive enough to reveal differences.

3.6.4. Limitations and Future Work

We examined average physiological responses over a play session, and did not test phasic responses to specific game events. Our results confirm that player experience was not different overall, but cannot confirm whether player responses within a play session were similar. It is possible that the two valence groups responded differently to specific game events (especially those that were targeted in the game reviews), but that these differences were not large enough in magnitude or frequency to affect average experience. Future work examining phasic responses to in-game events would help to shed light on the player reactions that make up overall experience. We gathered subjective reactions through game quality ratings. It is possible that although we saw no objective differences in play experience, the interpretation of similar play experiences in the context of different review text might cause players to subjectively characterize their play experience differently. Although outside the scope of this work, it would be interesting to know whether the cognitive rationalization of play experience with review text resulting in differing game ratings would also be reflected in differing subjective characterization of play experience.

Our study is the first use of physiological measures to evaluate experience prior to the playing of a game, which is important when considering biasing effects that can influence experience. The importance of pre-play events should not be dismissed, as we show that the

experience begins before engaging in actual play. This introduces a variety of interesting research questions as events prior to play deserve consideration, and more study is required to determine their impact on player experience. Our study only looks at biasing effects of game review text, but our methods could prove useful in the evaluation of any number of pre-play activities, such as effects of any pre-game menu traversal, loading times, or profile creation.

3.7. Conclusion

In this paper we have demonstrated the biasing effect that reading video game reviews has on player perceptions of a game. We show that these effects cannot be entirely explained by mediating changes in mood, and that the effects do not influence player experience as measured objectively through physiological sensors. Our major contribution is showing that the differences in perception of game quality that result from playing a game after reading a review are not stemming from differences in the play experience, but are a cognitive rationalization of the experience with the content and tone of the review. Our results are important for understanding player experience and for the game industry where reviews and commenting forums affect sales.

3.8. Acknowledgments

We thank NSERC and the GRAND NCE for funding.

CHAPTER 4

CRITIC-PROOFING: HOW USING CRITIC REVIEWS AND GAME GENRES CAN REFINE HEURISTIC EVALUATIONS

4.1. Overview

In Chapters 2 & 3 we explored the effects that game reviews and user comments have on a player's experience and their perception of game quality. The exploration gave us insight and understanding into the important effects that reading affective text can have on a player. Through our understanding it becomes clear that there is a need to strive to receive good reviews, not necessarily because doing so will improve the player's experience, but rather to avoid the feedback loop caused by negative reviews and comments.

In this chapter we switch gears from expanding our understanding of the effects of game reviews to using them to improve our usability evaluation methods. Heuristic evaluation – a technique where experts inspect software and determine where the application violates predetermined policies for good usability – is an effective technique for evaluating productivity software. In this chapter we describe a modified heuristic evaluation technique – called critic-proofing – that uses data mined from game reviews to create an ordered list of problems identified during the heuristic evaluation process, and we outline in future work in chapter 5 a potential system design that utilizes data-mining and machine learning to improve our critic-proofing technique.

We begin, in this chapter, with a description of our critic-proofing technique, and present a case study where we used our technique to evaluate a game in development. We found that the critic-proofing technique provided the developers with valuable additional insight over existing heuristic evaluation methods while adding negligible additional effort to the process.

4.2. Introduction

Development in the computer game industry is fast-paced and tightly scheduled. Iterative evaluation of games during development must adhere to this rapid schedule, making time-consuming approaches undesirable. Discount usability evaluation methods have recently been adapted for the games industry due to their low cost, short time requirements, and reliable feedback to developers. For example, in heuristic evaluation, evaluators examine features and system actions that violate predefined heuristics, and assign a severity rating to any discovered problem (Nielsen 2007) – a fast, cheap, and reliable method of evaluation. However, the interaction and interface components of computer games are different from productivity software, and require improvements to discount evaluation processes, such as heuristic evaluation, if they are to be relevant and useful to game developers.¹

Heuristic principles are often designed to generalize across applications, failing to take into account the diversity in game types, game controls, and game interfaces. The diversity in video games is a testament to creativity, but also drives variety in usability requirements so a given heuristic may not be equally relevant to all games. Despite the differences between games, reoccurring similarities between games exist within the same game genre. Often games that share a genre will have similar control schemes and interface layouts, for example, in the PC shooter genre W, A, S, and D are almost exclusively used for character movement, and the SPACE bar is used for jumping. Players who are accustomed to the shooter genre expect these control mappings. These genre characteristics can aid game designers by providing a template for the interaction, allowing new designs to borrow from previous games. While heuristic evaluation for

¹ This chapter contains sections of (Livingston, Mandryk, et al. 2010) that are reprinted here with the kind permission of the ACM

games must encompass a variety of different interfaces, taking the game's genre into account can produce a superior starting reference point for analysis.

A heuristic evaluation results in a list of usability problems and an accompanying severity indicator. These severity ratings are generally grouped into coarse-grained categories based on the problem's impact, frequency, and persistence, and are determined subjectively by an expert evaluator. Although a particular problem can be classified as severe, there is no method for prioritizing the problems within a category. In addition, there is no means of prioritizing problems based on the game's genre. For example, poor control mappings will be a more significant usability issue for an action game than a strategy game (Pinelle, Wong & Stach 2008a). Our method addresses the issue of genre specifically by using critical reviews as a method of determining the importance of a usability problem. The more often critics encounter a specific usability problem within games of the same game genre, the more important we consider that category of usability problem.

By incorporating critical reviews into our evaluation process we move towards an approach that will 'critic-proof' games. The goal of a game developer is to release a game where all potential criticisms by reviewers have already been identified and fixed. Critic-proofing a video game is a difficult task, especially when a game is entirely new and there are no reviews of previous versions. Our modified heuristic evaluation approach describes a method for developers to use in their evaluation process to critic proof the usability of their games prior to release. For this paper we refer to the heuristic evaluation process described by Nielsen as the traditional heuristic evaluation process.

We modified the traditional heuristic evaluation process to address the following problems:

1. The feedback from heuristic evaluations tends to be grouped into coarsely-grained categories instead of prioritized lists.
2. Any prioritization that does result is a subjective choice.
3. Heuristic evaluation fails to take the diversity of video games and game genres into account.

We employed our evaluation technique in a case study of a video game currently in development. The results from the case study suggest that our technique provides significant value in the identification and prioritization of usability problems.

In this paper, we present relevant prior work on the heuristic evaluation process, video game genres, and the concept of ‘critic-proofing’ a game. The evaluation technique is then presented in detail, describing how an evaluator would generate a prioritized list of usability problems. Next, a case study is presented, where we analyze the use of our evaluation technique on a game currently under development in a natural development setting. Finally, we discuss the advantages and limitations of our technique and our future plans for research in this area.

4.3. Related Work

Previous research has investigated the importance of game heuristic techniques for a conceptual understanding of playability (Nacke 2009). In our work, we have furthermore identified four relevant areas of work that directly relate to the heuristic technique used in this paper: (1) severity ratings, (2) video game heuristics, (3) critic proofing, and (4) video game genres.

4.3.1. The Heuristic Evaluation Technique

Heuristic Evaluation is one of the discount methods used for usability inspection (Nielsen & Mack 1994). In heuristic evaluation, human-computer interaction experts inspect a user interface and identify usability problems. Usability problems are categorized into ten heuristics

defined by Nielsen (Nielsen 2005), which focus on issues like providing help and consistency. In addition, usability problems are given a severity rating based on the impact, frequency, and persistence of the problem. Heuristic Evaluation is one of the most commonly used software usability evaluation methods as it is cheap, fast, and reliable, and can be conducted throughout the development cycle, on low-fidelity paper prototypes to fully-realized applications.

4.3.1.1. Video Game Heuristics

Heuristics are not new to game research. Federoff (2002) was one of the first pioneers into game heuristics. Through observations and conversations at a San Francisco game company, and a comprehensive review of literature, Federoff was able to compile a list of heuristics, which focus on three game areas: interface, mechanics, and play. Federoff's heuristics are quite broad, while we are specifically interested in problems pertaining to usability.

In 2004 Desurvire et al. (2004) developed a set of heuristics to evaluate game playability, or HEP. Although similar to Federoff's heuristics, Desurvire provided a simpler description and organization of the heuristics. Recently Desurvire has published a newer version of the HEP heuristics called PLAY (Desurvire & Wiberg 2009). In this she attempts to make the underlying principles more generalizable by applying the heuristics in three different genres during evaluations. In the scope of our paper the HEP and PLAY heuristics apply to far more than just usability, and the genre differences were only briefly explored.

Other forms of heuristics have been developed to fill different needs of game evaluation. Korhonen et al. created heuristics for mobile multiplayer games (Korhonen & Koivisto 2007); Pinelle et al. generalized this idea with a set of heuristics for all multiplayer games (Pinelle et al. 2009). Many of these heuristics have similarities and overlap, and some attempts have been made to compare techniques. Korhonen et al. compared two different sets of playability

heuristics – HEP (Desurvire et al. 2004) and their own mobile games heuristics (Korhonen & Koivisto 2006) – to examine the relative strengths and weaknesses of each approach (Korhonen et al. 2009). Koeffel et al. (2010) attempted to create a comprehensive list of heuristics from those found in literature, and assess the collection’s effectiveness using a comparison to video game reviews.

In our study we have employed the usability game heuristics created by Pinelle et al. (Pinelle et al. 2009). While others have also provided usability heuristics, Pinelle’s list is specific and short—there are ten principles—making the evaluations clean and simple to perform. Also, in addition to this, Pinelle’s work has been extended in the exploration of usability specific to genres (Pinelle, Wong & Stach 2008b).

4.3.1.2. Severity Ratings

The method of categorizing usability problems is key to deciding which should receive developer attention first. Most recent work in the area of heuristics research focuses on the development or modification of the heuristics themselves, and on testing their validity. These game-specific heuristic techniques use one of two methods for assigning severity ratings, both developed for evaluating productivity applications: Nielsen’s severity rating (Nielsen 1995), or Dumas and Redish’s severity rating (Dumas & Redish 1993).

Both methods (Nielsen’s and Dumas’s) rely on a combination of evaluator and development judgment in determining the severity of any discovered problems (Nielsen 1992). Nielsen describes a five category method ranging from “*I don't agree that this is a usability problem at all*” to “*Usability catastrophe: imperative to fix this before product can be released*” with the severity being assigned based on the frequency, impact, and persistence of a problem (Nielsen 1995). Dumas and Redish, on the other hand, use four levels ranging from, “*Level 1*

problems prevent completion of a task.” to “Level 4 problems are more subtle and often point to an enhancement that can be added in the future”. In addition Dumas assigns one of two scopes to a problem, either: local—applying in a single instance such as one screen, or dialog box—or global—applying to more than one instance perhaps across multiple screens, or systemic across all dialog boxes (Dumas & Redish 1993). Both approaches are limited in two ways: they rely solely on human appraisal, and are coarse in their granularity. Together, these limitations mean that many problems are simply assigned the same severity rating.

Recently there have been attempts to map heuristic results to the ratings of game critics. In a recent book chapter, Koeffel et al. (2010) presented a study where heuristic evaluations were performed, using Nielsen’s ratings (Nielsen 1995), and a score was assigned to each problem based on its severity. The score was then compared to scores received from critics. This is a valuable step forward, however Koeffel et al. still rely on a purely human appraisal method, and while they discuss the importance of game genres, they don’t account for usability differences found between genres.

4.3.2. Critic-Proofing

This paper describes a technique that uses the reports of game critics to address common usability problems before post-release critical analysis exposes them. By identifying interface factors that have historically caught critics’ negative attention, we can prioritize effort to minimize historically common shortcomings.

The voice of a professional critic is powerful. A single influential critic can have a significant effect on any product, influencing consumers purchasing decisions. Larsen (Larsen 2008) explored how a critic’s review is in essence an unscientific user experience evaluation. Larsen identifies that game reviewers generally provide both a qualitative (textual review) and

quantitative (score) component. Some reviewers evaluate different aspects of a game separately. In many cases these aspects are comparable to hedonic and pragmatic aspects of user experience research.

Koeffel et al. also offers support for the use of the reviews written by game critics (Koeffel et al. 2010). In their recent study, Koeffel et al. compare results of heuristic evaluations to the average critic scores of games, and suggest that a similar trend between review scores and the number and severity of usability problems exists.

The term critic-proofing has recently been used to describe the technique used at BioWare in MassEffect 2. In an interview with Gamasutra (Nutt 2010), producer Adrien Cho described how the development team mapped critic and player feedback from the first game onto the design objectives for the sequel.

“But the other part of the equation was actually taking all the feedback -- I'm not saying some -- absolutely every feedback from press and the fans, and collating all that into a huge list.

Everything eventually fit within certain categories, and when we looked at that, mapped with the things that we wanted to fix, it became really clear. It became a blueprint. It made making the sequel really easy.” (Nutt 2010)

This critic-proofing process can only be applied to a sequel, as critic feedback is difficult and expensive to collect before release.

4.3.3. Video Game Genres

Our work draws heavily from work published by Pinelle et al. in (Pinelle, Wong & Stach 2008a) and (Pinelle, Wong & Stach 2008b). The underlying premise of Pinelle et al.'s work, as established in (Pinelle, Wong & Stach 2008a), is that game reviews can be used as a proxy for providing empirical data for deriving game heuristics when a substantial body of scientific

research is not available. To support this hypothesis, they analyzed 108 reviews by 24 different reviewers of moderately to poorly rated games from a popular game review website, and derived a set of novel heuristics for game interface issues, summarized in Table 4.1. They tested these heuristics by having experts use them in a test of a different, but completed game not included in the original review set.

In their subsequent publication (Pinelle, Wong & Stach 2008b), Pinelle et al. further probed their dataset to determine if usability problems from the reviews were correlated with game genre. They used the genre definitions from the review site, which was composed of the common Role playing, Sports, Shooter, Action, Strategy and Adventure genre classifications. Statistical tests were used to show that while some problems spanned all genres, many were genre specific; for example “Skip Content” was only a problem in Adventure games, while controls issues were critical in Action and Sports genres.

Table 4.1 Pinelle’s 12 usability problems found in video games.

Problem category	Key issues
1. Consistency	poor hit detection, poor in-game physics, inconsistent response to input
2. Customizability	does not allow user to change video and audio settings, difficulty, or game speed
3. Artificial intelligence	problems with pathfinding, problems with computer controlled teammates
4. View mismatch	bad camera angle, view is obstructed, view does not adjust to user’s action quickly enough
5. Skip content	cannot skip video and audio clips, frequently repeated sequences
6. Input mappings	bad input mappings, limited device support, limited control customization
7. Controls	oversensitive controls, unnatural controls, unresponsive controls
8. Game status	does not provide adequate information on character, game world, or enemies. visual indicators, icons, and maps are inadequate.
9. Training and help	does not provide default and recommended choices; does not provide suggestions and help; does not provide adequate documentation, instructions, tutorials, and training missions.
10. Command sequences	learning curve is too steep; requires too much micromanagement; command sequences are complex, lengthy, and awkward, making the game difficult to play
11. Visual representations	bad visualization of information, too much screen clutter, too many characters or game elements on the screen at the same time, difficult to visually distinguish interactive content from non-interactive content
12. Response times	slow response time interferes with user’s ability to interact with the game successfully

While Pinelle et al. discussed the implications for design in (Pinelle, Wong & Stach 2008b), they did not provide any empirical evidence for the utility of their genre-specific evaluation technique. They also noted several potential opportunities and shortcomings: first, that many problems are of moderate impact and could be ignored if developers focused only on high-impact issues; second, that genre is a fluid idea, and that many games span genres or contain mini-games or components from several different genres; and finally that while their findings highlight areas of specific interest by genre, they do not provide a mechanism for compromising between broad and focused development.

Our approach codifies a heuristic technique based on the genre and overall rating systems of Pinelle et al. (Pinelle, Wong & Stach 2008b), while addressing the shortcomings they identified to provide a simple and robust method for game usability analysis and critic-proofing.

4.4. Critic-Proofing Heuristic Technique

When compared with the heuristic approach presented by Nielsen (Nielsen 2007), our evaluation system requires little additional complexity while offering substantial additional value. The two techniques are comparable for the evaluation and determination of severity classification. Our approach adds two additional steps – pre-evaluation to identify the appropriate severity framework, and severity calculation to prioritize the usability findings.

4.4.1. Technique Overview

Our technique is a simple process that can be added to a traditional heuristic evaluation of a game at any point in the evaluation process. The additional steps we describe are best utilized in the manner we outline—the pre-evaluation at the beginning, and the severity calculation after problem categorization—however the technique is flexible enough that our process may be applied to a completed evaluation.

To perform our technique the steps in the process are:

1. Determine game and sub-component genre ratings.
2. Evaluated entire software for heuristic principal violations.
3. Subjectively assign classification ratings based on the impact, frequency, and persistence of the violation.
4. Calculate the prioritized severity rating according to problem classification, game genre, and the heuristic violated.
5. Average prioritized severity ratings across all evaluators.

The technique adds little additional overhead to the overall evaluation process, as steps 2 and 3, which are common to all heuristic evaluation techniques, dominate the effort. Steps 1 and 4 may be performed post-hoc, but as heuristic evaluation should be completed individually by each evaluator, averaging across evaluators should always be the last step performed.

4.4.2. Pre-evaluation

The pre-evaluation step should be performed prior to inspection, to establish the severity frameworks used to prioritize problems.

4.4.2.1. Assessing Gameplay & Determining Genres

Our first step assesses both the game and gameplay elements to be evaluated. We define a gameplay element as a subsystem of a game that can be described as a separate and complete play experience, particularly identifying mini-games separately for evaluation. This assessment is required to ensure the genres for the game and gameplay elements can be identified.

Every game can be categorized into one or more genres. Our technique uses the six primary genres presented by Pinelle et al. (2008b). In their work, each genre has a rating for its applicability to each heuristic category – a genre rating. Rather than assuming that each game

falls solely into a single genre, we calculate genre ratings using the proportion of the game that falls into each genre.

4.4.2.2. Calculate Genre Rating

Because each genre in Pinelle et al.'s (2008b) work was found to have distinct usability trends, it is reasonable to base the genre ratings on their observations. We use the trends and the values they observed for each problem category in each genre as the basis for calculating a normalized genre rating. If the game or game elements fall completely into a single genre, then the normalized ratings maybe used directly. However, if the game or elements fit into more than one genre, a weighted average must be calculated using formula 1:

$$R_{g_{norm}} = \frac{\sum_{i=1}^n w_i R_g}{g_{max} \sum_{i=1}^n w_i} \quad (1)$$

Where w is the genre weight (the proportion of the game that is identified as each genre), R_g is the genre rating, and g_{max} is the maximum possible value of R_g . In this paper the summation of w_i is 1; however, to ensure that formula 1 remains generalized, the summation denominator has been included. This step is completed for each heuristic category separately. Although this equation may appear to be complicated, the approach is straightforward. Consider a game that has been identified as 20% Action and 80% Shooter and we are interested in the Controls heuristic. Pinelle et al.'s ratings for Controls are 12 for Action and 5 for Shooter. The new genre rating would be:

$$R_{g_{norm}} = \frac{(0.2 * 12) + (0.8 * 5)}{12} = 0.533$$

In this paper we employ genre ratings taken from Pinelle et al. (2008b) summarized in Figure 4.1, but our system applies equally to other genre ratings frameworks or studies.

Once the normalized genre ratings have been calculated they can be used in the severity calculation step. This new step is performed following the heuristic inspection, and is used to produce a prioritized list of potential usability problems.

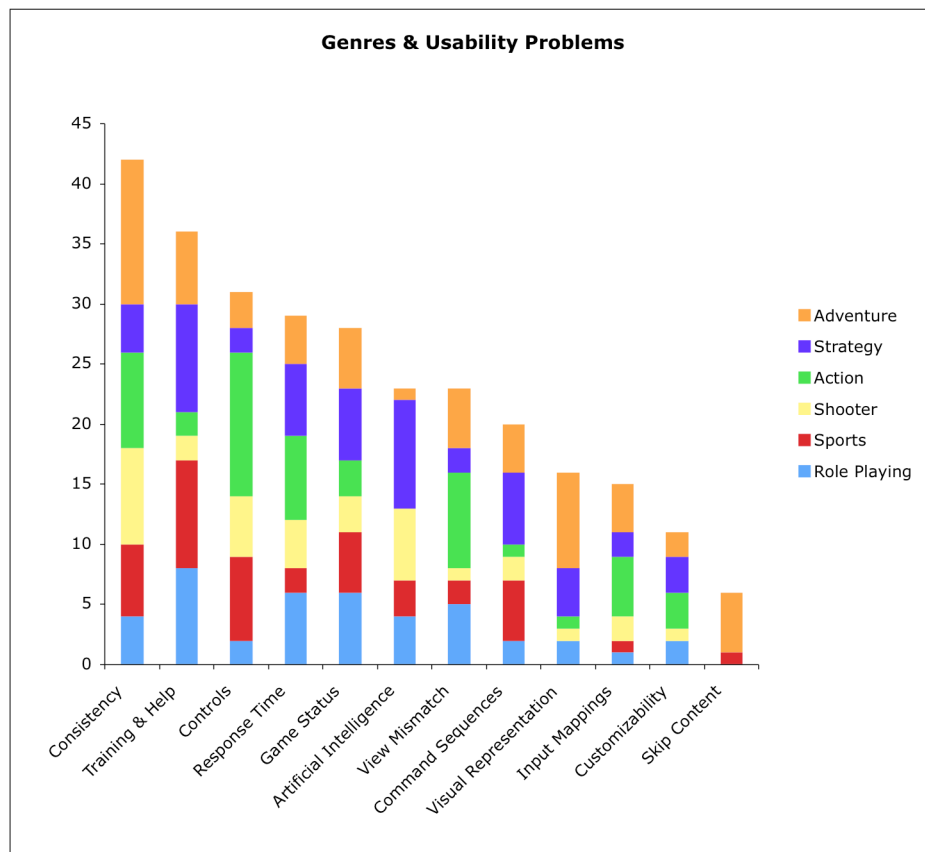


Figure 4.1 The 12 common usability problems combined across genre into frequency of occurrence of the problem.

4.4.3. Severity Calculation

In traditional heuristic evaluation, the severity rating is limited by the categorizations available. In many ways the severity rating is akin to triage, coarsely grouping problems into a relatively small number of categories. For example, Nielsen only has five ratings (Nielsen 1995), and in his book, Dumas (Dumas & Redish 1993) only list four ratings and two different scopes. For traditional software artifacts these categories are sufficient. However, the fast pace and aggressive development cycles in the game industry reduce the value of the evaluation, because a problem's importance within a specific category is ambiguous. It is reasonable to imagine that two problems in a game may both be categorized as a major problem (using Nielsen's ratings) with one problem being significantly more important than the other. Our system removes this ambiguity by assigning a finer-grained severity rating to usability problems, based on empirical evidence presented for specific genres (Pinelle, Wong & Stach 2008b). Each heuristic violation must be categorized into one of the 12 problem categories found in Table 4.1. The genre rating for the categories can then be used in the prioritized severity calculation.

Table 4.2 Normalized Classification Ratings for a slightly modified Nielsen severity ratings.

Classification Rating	Severity Description
-1.00	Not A Problem
+0.00	Minor
+0.33	Moderate
+0.66	Major
+1.00	Catastrophic
+2.00	Game breaking

Our approach classifies problems using a modified set of the severity rating descriptions presented by Nielsen (Nielsen 1995). We add one additional rating, ‘*game breaking*’, to the set. We have assigned values, ranging from negative one to positive two, to each of these descriptions (Table 4.2). There are two special case entries in this table. The first is the ‘not a problem’ description which has been assigned a classification of -1. This description is used for a problem that has been mistakenly identified. By assigning a value of negative one we ensure that any problems classified in this way will have a negative final normalized priority rating. This will be the case regardless of the genre rating, effectively removing those problems from the list. Similarly the new ‘*game breaking*’ description refers to a usability problem that renders the game unplayable. Any problem with this classification will be assigned a prioritized rating greater than one, implying it should be addressed regardless of the genre rating.

Any numerically-different classification system may be used in principal, as long as it is normalized according to formula 2:

$$R_{c_{norm}} = \frac{R_c}{c_{max}} \quad (2)$$

Where R_c is a numerical classification rating, c_{max} is the maximum value of R_c , and $R_{c_{norm}}$ is the normalized numerical classification rating.

We then perform a calculation to determine the prioritized severity rating using formula 3:

$$R_p = \frac{R_{g_{norm}} + R_{c_{norm}}}{2} \quad (3)$$

Where R_{gnorm} is the normalized genre score ranging from 0 to 1, R_{cnorm} is the normalized rating assigned to heuristic severity descriptions, ranging from -1 to 2, and R_p is the final prioritized rating for the usability problem, which ranges between -1 and 2. Generally ratings calculated this way will generally range between 0 and 1, with ratings below 0 or above 1 warranting special treatment. By providing finer granularity in the severity ratings, we remove the guesswork of assessing priority for identified usability problems, providing greater confidence in the results of the evaluation.

The advantage of normalizing all ratings throughout this process is that both the genre rating and classification rating systems can be modified while maintaining a consistent prioritization system.

As an example lets imagine we have an action game that violates Pinelle's 7th game usability heuristic (*"Provide controls that are easy to manage, and that have an appropriate level of sensitivity and responsiveness."*), and the problem has been classified as *minor* (0.00). Lets also imagine there is a problem that violates Pinelle's 9th heuristic (*Provide instructions, training, and help.*), which is classified as a *major* (0.66) problem. The first problem is a *Controls Problem* with a rating of 1.00, and the second is a problem with *Training & Help*, which has a genre rating of 0.17. In this case the calculations would be as follows:

$$R_{controls} = \frac{1.00 + 0.00}{2} = 0.500$$

$$R_{help} = \frac{0.17 + 0.66}{2} = 0.422$$

As can be clearly seen, even though the usability problem relating to training and help was classified as a major usability problem, the minor usability problem is assigned a higher priority rating, because of the consistent negative impact of controls problems in the action genre (see Figure 4.1).

In this section we presented a novel critic-proofing technique for usability evaluation. Our technique differs from previous heuristic evaluation techniques in the unique way that it incorporates common genre problems into the severity calculations. We presented two additional simple steps that, when added, will provide developers with a prioritized list of usability problems. The following section presents a case study where our technique was applied successful to a game currently under development.

4.5. Case Study: Capsized

In this section we present a case study examining the use of our heuristic evaluation technique on a game currently under development. We will cover the original evaluation of the game using a traditional heuristic process with the heuristic principles presented in (Nielsen 2007) and (Nielsen 2005). We will discuss the implementation of our modified evaluation technique and its application to usability problems identified in the original heuristic evaluation. We will also present a comparison of the two techniques highlighting differences in the evaluation findings. We conclude this section with a discussion of feedback from the developer.

4.5.1. The Game

Capsized (Figure 4.2) is a PC and xBox Live Arcade (XBLA) game in development from Alientrap Software, designed by Lee Vermeulen (programmer) and Jesse McGibney (artist), using sound effects by Michael Quinn and music (used with permission) from Solar Fields.

The gameplay is based on classic platformers, focusing on taking the intensity of the first person shooter genre into the 2D world. The gameplay is enhanced with the use of rigid body physics to allow for more player interaction with the environment, giving the user a unique experience while exploring the world. In the story, a space traveler whose ship has 'capsized' on an unknown planet must avoid hostile alien inhabitants and rescue his stranded crewmates. To win, the player must finish 15 levels, all with different objectives, puzzles, enemies, and weapons.



Figure 4.2 Gameplay screenshot from Capsized

To interact with the game, players use both the keyboard and mouse. Mimicking the traditional shooter control scheme, the mouse is used for aiming and firing and the keyboard is used for movement. Although the game can be played using a standard xBox controller, the evaluation presented here was performed on the PC version of the game. The game also features

a traditional non-diegetic HUD system displaying the player's current health, jet fuel, objective direction, and ammunition.

4.5.2. A Basic Evaluation

Before the heuristic evaluation was conducted, Pinelle et al.'s heuristic principals were presented on the Capsized Play Testing & Development Forum. The beta version of the game was tested, representing a vertical slice of the entire game.

After the initial evaluation was conducted, one of the six (Table 4.2) severity classifications were assigned to each of the 28 identified usability problems. The problems ranged in severity from *minor* to *major*. No *Catastrophic* or *Game Breaking* problems were found. The majority of problems were classified as *moderate* or *minor* with a small portion being classified as *major*, and one classified as *not a problem* (Figure 4.3).

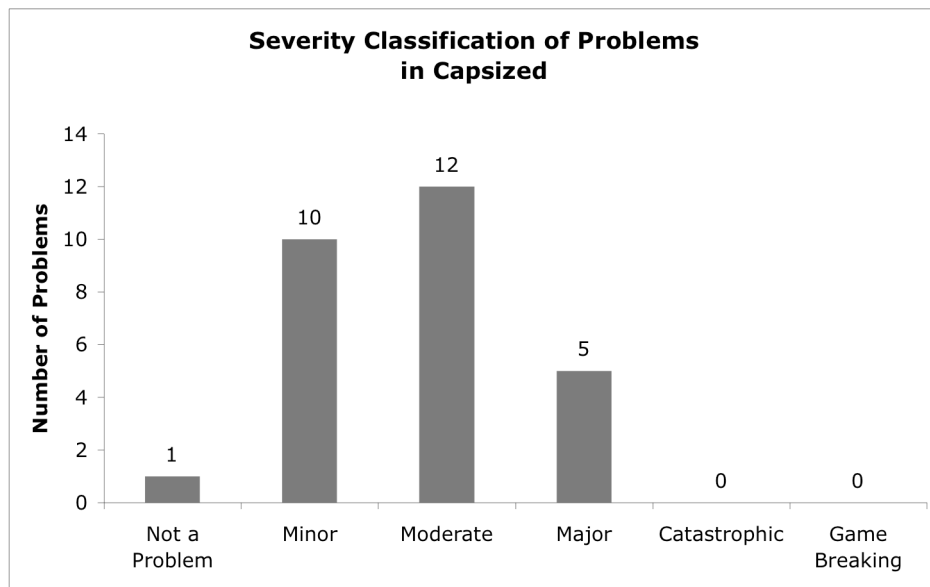


Figure 4.3 This graph shows the distribution of usability problems identified in Capsized, October 2009

The initial evaluation was sufficient to identify the usability problems; however, the coarse severity ratings were insufficient for the needs of the developer. Thus, we applied our new critic-proofing technique to the evaluation results in a post-hoc manner.

4.5.3. Critic-Proofing Heuristic Approach

Before the new severity ratings could be applied to the usability problems, Capsized needed to be assessed using our pre-evaluation step. Based on feedback received from the developer, the game was classified into the action and shooter genres. Each genre was weighted at fifty percent and a weighted average was calculated providing the genre ratings for the game (Table 4.3).

Table 4.3 The calculated genre ratings for Capsized based on the Action and Shooter Genres.

	Action	Shooter	Genre Rating
Consistency	8	8	0.667
Training & Help	2	2	0.167
Controls	12	5	0.708
Response Time	7	4	0.458
Game Status	3	3	0.250
Artificial Intelligence	0	6	0.250
View Mismatch	8	1	0.375
Command Sequences	1	2	0.125
Visual Representation	1	1	0.083
Input Mappings	5	2	0.292
Customizability	3	1	0.167
Skip Content	0	0	0.000

Each of the previously identified heuristic problems was assigned to one of the 12 common usability problems, and we performed our new severity rating calculation using both the newly-calculated normalized genre ratings, and the normalized classification ratings. The technique produced prioritized ratings ranging from 0.684 to -0.167, with 18 unique classifications.

The increased granularity of severity made the organization of usability problems more obvious and provided the Capsized developers with a report outlining the most significant usability problems. Both heuristic evaluation results were presented, and feedback comparing the two techniques was collected.

4.5.4. Comparing Results

The initial severity classifications were done with text descriptions, (i.e. minor, moderate, major). To compare these descriptions directly with the prioritized ratings calculated in our technique we used the normalized numerical classification ratings found in Table 4.2.

We carried out a direct comparison between the results of the two techniques. In Figure 4.4, we present the normalized severity ratings for both techniques. The figure shows the severity ratings ordered by the prioritized rating in decreasing severity with the corresponding classification rating and genre rating that a specific problem received. The chart shows that some usability problems are reordered when genre and classification ratings are combined into the prioritized rating, creating a much clearer rating system.

Here are two examples of usability problems from the traditional evaluation:

- *Description:* Grappling hook rope length uncontrollable

Heuristic violated: 7 Provide controls that are easy to manage, and that have an appropriate level of sensitivity and responsiveness.

Severity: **moderate** (The problem is persistent and occurs frequently, but has a low impact because it can be worked around)

- *Description:* No option to skip tutorial or story text if you desire, info bubbles are difficult to avoid and cannot be sped up.

Heuristic violated: 5 Allow users to skip non-playable and frequently-repeated content.

*Severity: **major*** (This has a high impact, frequency, and persistence, especially when playing a level more than once.)

In this case the first problem is grouped with 11 other moderate problems and the second with 4 other major problems. Also the second problem is considered more severe than the first. If we examine the same problems evaluated using our technique:

- *Description:* Grappling hook rope length uncontrollable

Heuristic violated: 7. Provide controls that are easy to manage, and that have an appropriate level of sensitivity and responsiveness.

*Severity: **0.519 Controls moderate*** (The problem is persistent and occurs frequently, but has a low impact because it can be worked around)

- *Description:* No option to skip tutorial or story text if you desire, info bubbles are difficult to avoid and cannot be sped up.

Heuristic violated: 5 Allow users to skip non-playable and frequently repeated content.

*Severity: **0.330 Skip Content major*** (This has a high impact, frequency, and persistence, especially when playing a level more than once.)

Using our technique the moderate controls problem is rated as more severe because in game reviews controls were cited as a problem more often than skip content for action/shooters.

The results of the evaluation were presented to the Capsized development team in a report. Both the classification ratings and the prioritized ratings were presented in the report; however, the prioritized ratings were assigned a numerical value, while the classification ratings were presented in text. No problems were severe enough to be classified as either ‘*game breaking*’ or ‘*catastrophic*’. The 28 problems were grouped into one of the other four categories: five were ‘*major*’, twelve were ‘*moderate*’, ten were ‘*minor*’, and one was ‘*not a problem*’.

Problems in the report were ordered from most severe to least severe. The second report contained a list of the same problems; however, they were prioritized using our technique. The prioritized severity ratings were multiplied by 100 to ensure ease of reading, but this has no bearing on the function of the rating system.

4.5.5. A Developers Reception

Play testing began in October 2009. This build represented 1/3 of the final game, and included most of the game mechanics. Before the evaluation was performed, the heuristic evaluation technique was described and presented to the designers. In response, one designer said, “The list of usability principles was a great read, I look forward to seeing your full review.”

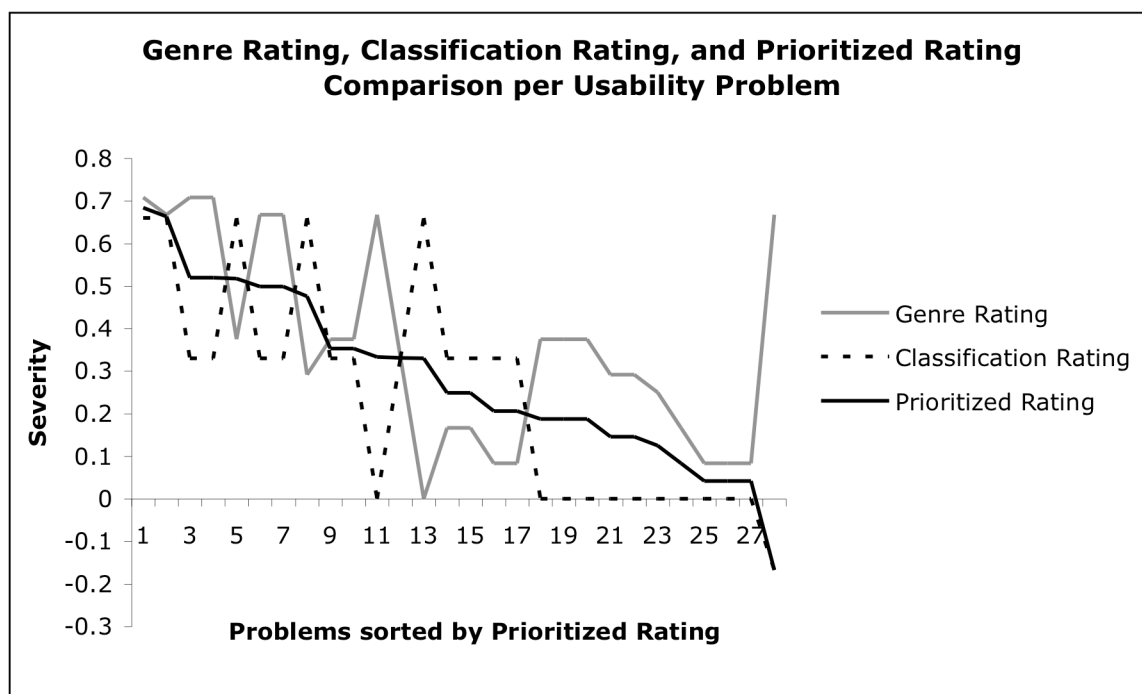


Figure 4.4 A comparison between severity ratings demonstrating the advantage of the prioritized rating.

The designer was questioned after he had examined both the traditional and critic-proofed heuristic analysis. When ask which of the two was more valuable he picked our prioritized list:

“I would say the priority list is more important.”

And:

“It's more developer oriented and task based, and can be constantly updated and items moved around based on what needs priority.”

It was also clear that heuristic evaluation in general was considered to be something valuable that multiple developers could perform, thereby increasing the visibility of usability issues:

“It would be constantly updated and more developers would take notice of these issues.”

And:

“With this development-centric approach we will be able to solve these problems more effectively, and clearly identify more possible issues.”

We also drew a comparison to more traditional playtesting models (Fullerton et al. 2008). The designer clearly felt that our technique provided greater benefit compared with traditional playtesting. Suggesting there would be value in allowing playtesters to perform the evaluations:

“Rather than just a description of what was 'fun' or 'exciting' (which is of course very valuable), the system clearly identified problems and their severity, making it more development-centric than the evaluation system in place.”

And:

“I would most likely base it on a wiki and host the priority system there once production stage is done and we are focusing on playtesting. So even playtesters could add to the system.”

4.6. Discussion

The technique we have outlined is a valuable addition to the game-testing process. Our approach incorporates the advantages of traditional heuristic usability evaluation, such as cost effectiveness and small time investment, while also providing designers with a new empirically-

grounded method to evaluate the severity of problems. Many inspection methods rely heavily on designer judgment and experience for prioritizing problems; our technique allows another avenue for validation. We do this by providing a metric-based approach to help determine where developer resources should be focused, based on critic reviews from prior games.

Play testing can be performed on a game as soon as it reaches a playable state. This testing is typically done in-house, or out sourced to a dedicated testing company. Actual players don't test a game until it reaches a later stage of development, often because developers don't want their game becoming public too early. If a game becomes available too early, players' and critics' perceptions about quality and content might be biased before the final product is released.

Although useful for companies of all size, our technique is particularly valuable for a small development team. The process is simple, and heuristic evaluations are quick and cheap. In addition, the organizing and prioritizing of problems that our technique provides helps to mitigate the development risk of limited corporate product memory for new independent studios and inexperience.

Larger companies and publishers can also benefit from our technique. Our process can easily be added to any existing testing or development framework. By presenting heuristic feedback in an organized and prioritized manner, we enable larger teams to integrate heuristic testing more tightly with project management systems and tools. Larger development teams can also benefit from the process in many of the same ways that smaller teams do – by performing usability tests cheaply and quickly at many points in the development process. The outsourcing of quality assurance (QA) is a common practice in established studios with publishing partnerships. QA is often performed by the publisher, and in many cases outsourced again. Our

technique also fits well in this model, because, prior to being delivered to QA, our technique can quickly be performed by the developers.

4.7. Conclusion

We have presented a modification to traditional heuristic evaluation of games that does not require much effort, time or resources, but provides substantial additional value by creating a genre-specific prioritized list of usability problems. When you compare our technique to traditional play testing, the value becomes apparent. Our technique can be performed early in a project's development, is fast and cheap, and provides developer-centric results. Our technique also incorporates historic genre-specific usability issues collected from game reviews. The evaluation technique may be performed prior to release, but provides feedback based on similar games.

We validated our technique with a case study of a game currently in production. Feedback from developers suggests that the technique provides additional benefit over both traditional heuristic evaluations and standard play testing.

We describe our process as critic-proofing, because the genre ratings are grounded in an empirical analysis of critical reviews. As we described earlier, feedback from critics is extremely valuable in assessing user experience (Koeffel et al. 2010; Larsen 2008). While the technique does not attempt to inform the designer of how to make a perfect game in the eyes of critics, it does incorporate critics' feedback to produce priority ratings of usability problems that can help focus development efforts.

4.8. Future Work

The technique that we have outlined is an important first step on the road to developing a more robust overall game testing approach based on user experience. Future studies will examine

the use of heuristic evaluations for the validation of system architecture, and the integration of the technique into a higher-level design methodology, such as scrum. To further validate our technique, we will conduct an empirical study comparing our results to critics' reviews. Our study explored a limited set of heuristics focusing exclusively on usability. Future studies will explore the use of our technique with other heuristics exploring topics such as playability.

4.9. Acknowledgments

We would like to say a big thank you to Lee Vermeulen who provided us with substantial feedback, even during development crunch time on the Capsized project. Without his input our evaluation would not have been possible, or nearly as successful.

CHAPTER 5

CONCLUSION & FUTURE WORK

5.1. Conclusion

Wherever media forms exist, we can be sure that there will also exist the voice of a critic. The critic plays an important role in our understanding both of the material itself and the cultural effects of that media. Video games are still a very young form of media, and have only gained mainstream interest in the last decade. As such, there has been little study of the effects of reviews on the players of games. In this thesis we have presented work outlining the effects that game reviews have on user opinion and experience, and how we can use game reviews as a tool for improving usability evaluation methods. Our analysis provides a significant contribution to both our understanding of and our ability to leverage video game reviews. The importance of game reviews should not be understated – the influence of reviews and user comments can create a startling difference in how we perceive the quality of a game.

In chapter 2 we explored the effects of user comments and game reviews on player perception of experience. We showed the simple act of reading affective text that relates to the game to be played is enough to impact the players' perceptions of game quality. Through our controlling of factors we were able to determine that the effects were not solely mediated by changes in the players' moods as a results of reading the affective text. Interestingly, the effects were only present when participants read the negative texts. The results demonstrate the impact that reviews and other users can have on a player's perception of experience. While the results may appear to be obvious, they also raise a number of important questions about the possible impact of other forms of pre-play biasing factors, such as marketing.

Chapter 2 contributes to our understandings of the influence of video games. We provide 5 unique contributions: 1) we show that players who read the negative text rated the game significantly lower than players who read positive text. 2) The reviews or comment text had to relate directly to the game to produce biasing effects. 3) The authority of the source did not create different biasing effects. 4) The effects could not be solely attributed to changes in mood as a result of reading the text, as measured using self-report. 5) Negative text had a larger biasing effect than positive text.

Chapter 3 extends our work from chapter 2. We explore the objective effect of game reviews on player experience, finding that while reviews consistently impact a player's perception of overall quality, their objective experience remains consistent and unchanged. We attribute this effect to a cognitive rationalization of the experience with the text that was read. We demonstrate that different players may *feel* the same about the game, but that they each *think* that the quality of the game was different based on the negative or positive text that they read. These results are interesting because of the questions that they raise about biasing factors in perceived game quality. Our study also opens the door to the evaluation of pre-play events; an area that has previously received little study, but contains many interesting questions to answer.

In chapter 3 we provide four important contributions extending our work from chapter 2. 1) We confirm our finding from chapter 2 where we observed that reading positive or negative game reviews influenced the players' subsequent ratings of game quality. 2) We demonstrate that the difference can't be explained by changes in player valence, but may have been mediated by player arousal. Increases in arousal after playing were significant for players who read positive text and only marginally significant for players who read negative reviews, causing a significant difference between the two groups of players. 3) We observed significant increases in

physiological measures when playing the game compared with reading of the reviews, which illustrates the sensitivity of the measures used. And 4) There were no differences between players' physiological measures in the different groups of participants. This suggests that the players' experience did not change as a result of reading the positive or negative reviews.

In chapter 4 we switched gears from the evaluation of the effect of game reviews and user comments on player experience to the practical use of game reviews in the evaluation of usability concerns. Our work in chapters 2 & 3 show us that the effect of a negative review or user comment can have a significant effect on a player's perception of a game, which can create a snowballing effect where negativity breeds negativity. To address the snowballing effect we presented a usability evaluation technique that accounts for problems identified in game reviews before a game reaches the market. The technique we presented is a new modified heuristic technique that used the contents of game reviews directly to evaluate the severity of usability problems identified during a heuristic evaluation. While we do not modify the heuristic technique directly, our method does produce a fine-grained prioritized list of usability problems. Our findings from our case study demonstrate the usefulness of game reviews in the context of usability evaluation techniques. We also show the importance of considering the diversity and differences found in games.

Chapter 4 provides four important contributions. 1) Our method provides a fine-grained prioritized categorization of usability problems. 2) We limit the subjectivity of the assignment of severity rating. We do this by calculating a portion of the rating using a formula based on the characteristics of the game. 3) Our technique shows that game reviews - divided by genre - can be used to produce a custom evaluation for any game. And 4) the critic-proofing technique demonstrates the effectiveness of using game reviews in an evaluation capacity.

Throughout each chapter we have demonstrated that effects and use of game reviews as a tool for improving game experience and designer insight. Despite the significant contribution of these works there remains a number of avenues of future work.

5.2. Future Work

Chapters 2 & 3 demonstrate the impact that a specific pre-play event (i.e., reading game reviews) can have on a player's perception of quality and on the player experience; however, there are many pre-play events that remain unexplored. The impact of menu navigation and loading screen duration might impact a player's perception of quality, especially if they play a role in player frustration. The assumption is that the longer it takes for a player to actually reach the point of play, the more frustrating the experience and, therefore, the lower the perception of quality. The critic-proofing method was designed to be extendable to other heuristic sets. However, we only evaluated the technique using Pinelle's game usability heuristics (Pinelle, Wong & Stach 2008a). One avenue of study would be to extend this technique to other heuristics to determine if the method is as extendable as we believe.

5.2.1. Automated System¹

Game reviews provide us with an excellent way of dealing with the limitations of current game heuristics. Unfortunately the current review-based heuristics (Pinelle, Wong & Stach 2008a) were limited and biased in their scope, focusing on only 6 game genres and 108 reviews from GameSpot (Pinelle, Wong & Stach 2008b). Also, the entire evaluation was conducted manually, with researchers visually searching the reviews for criticisms of usability problems. Since metacritic provides an aggregation of a large sample of game reviews, it is perfectly suited

¹ This section contains parts of (Livingston, Nacke, et al. 2010) that are reprinted here with the kind permission of NHTV Breda University of Applied Science

to provide a basis for improving and extending review-based heuristics. metacritic has also gained traction in the industry as a metric used by publishers to determine the success of a title and the potential return from a given game. We plan to use metacritic similarly to how Pinelle used GameSpot, to a) validate existing heuristics, b) use a tool already used by the games industry to provide understandable evaluation results, and c) demonstrate the potential value of using game specific evaluations.

5.2.2. Implementation Strategy

Examining game reviews manually is impractical on a larger scale. Instead, the process could be automated so a much larger review sample set can be utilized. To develop a more automated process we have developed a three-step implementation strategy (Figure 5.1).

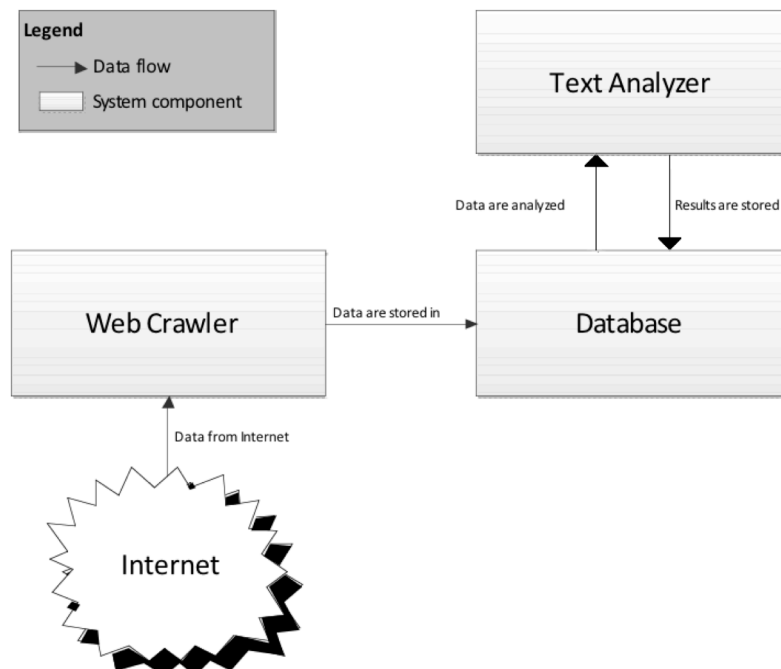


Figure 5.1 System Design

5.2.3. Web Crawler Functionality

Before any analysis can begin, review data must first be collected. This entails two types of data in addition to the more specific textual qualitative reviews and game scores (as well as

metascores). The web crawler would collect complete game reviews from the same websites used by Metacritic in the calculation of their ‘metascores’. The web crawler is an automated system that would run continuously, collecting new data as it’s posted and periodically rescanning to ensure that the data collected is up to date.

5.2.4. Data Collection Strategy

Metacritic is a rich source of data. In addition to the actual text reviews, we could also collect: platform, genre, release date, publisher, developer, number of players, rating, number of reviews, reviewer’s name, metascore, and any relevant or potentially important data from the specific game review sites we mine. We would store this data in a database for further analysis.

There exist two primary challenges that must be addressed in this data collection strategy. First, data must be collected from each review site that metacritic provides access too. However, each site provides different information formatted in different ways. The web crawler will need to be customized to ensure that it is capable of sorting data correctly when presented in different formats. This challenge is exacerbated because metacritic aggregates across more than one hundred game reviews. Second, metacritic uses some publications that are only available in print mediums, restricted to subscribers, or in a language other than English. The web crawler could not mine these publications; therefore we would have to omit them from our evaluation database. Fortunately, the number of restricted or print publications is small in comparison to those that remain, so a substantial number of reviews are still accessible.

5.2.5. Analysis Strategy

Perhaps the most challenging aspect, and one of the primary goals, for this project is the development of an automated analysis tool that would mine information from the text of game

reviews. Since game reviews are essentially untrained user experience reports, we could use the information from them to validate or reject game heuristics.

Analysis of the review data would be performed in a number of different ways. Initially the textual data could be analyzed using text analysis tools such as LIWC (Pennebaker et al. 2011) and ANEW (Bradley & Lang 1999). The results from LIWC would be examined to determine how words are used within a specific text. The use of heuristic-specific words could be used to determine the frequency of a specific heuristic problem, for example artificial intelligence issues. ANEW is a similar tool to LIWC except focusing specifically on the emotional content of words.

5.2.6. Final Thoughts

It remains clear that we have only just scratched the surface of the usefulness of game reviews as a tool for evaluation. However, the collection of works presented in this thesis has created a useful stepping-stone for the expansion of our general understanding of the significance held by game reviews. My current and future work will continue to explore this space, demonstrating the complexity, effects, and value of game reviews.

REFERENCES

- Baumeister, R.F. Bratslavsky, E., Finkenauer, C., Vohs, K.D., 2001. Bad is stronger than good. *Review of general psychology*, 5(4), p.323–370.
- Bradley, M.M. & Lang, P.J., 1999. *Affective norms for English words (ANEW): Instruction manual and affective ratings*, University of Florida: The Center for Research in Psychophysiology.
- Chaiken, S. & Eagly, A.H., 1989. Heuristic and Systematic Information Processing within and. *Unintended thought*, p.212.
- Chaiken, Shelly, 2003. Heuristic versus Systematic Information Processing and the Use of Source versus Message Cues in Persuasion. In *Social Psychology: A General Reader*. Psychology Press, pp. 461-473.
- Chaiken, Shelly & Maheswaran, D., 1994. Heuristic Processing Can Bias Systematic Processing: Effects of Source Credibility, Argument Ambiguity, and Task Importance on Attitude Judgment. *Journal of Personality and Social Psychology*, 66(3), pp.460-473.
- Desurvire, H. & Wiberg, C., 2009. Game usability heuristics (play) for evaluating and designing better games: The next iteration. *Online Communities and Social Computing*, pp.557–566.
- Desurvire, H., Caplan, M. & Toth, J.A., 2004. Using heuristics to evaluate the playability of games. In *CHI'04 extended abstracts on Human factors in computing systems*. p. 1512.
- Drachen, A., Nacke, L.E., Yannakakis, G., Pedersen, A., 2010. Correlation between heart rate, electrodermal activity and player experience in first-person shooter games. In *Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games*. Sandbox '10. New York, NY, USA: ACM, p. 49–54.
- Dumas, J.S. & Redish, J., 1993. *A practical guide to usability testing*, Ablex Pub. Corp.
- Ekman, P., 1994. Moods, Emotions, and Traits. In *The nature of emotion: fundamental questions*. Oxford University Press, pp. 56-58.
- Federoff, M.A., 2002. *Heuristics and usability guidelines for the creation and evaluation of fun in video games*. Citeseer.
- Fiske, S.T., 1980. Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, 38(6), pp.889-906.
- Fullerton, T., Swain, C. & Hoffman, S., 2008. *Game design workshop: a playcentric approach to creating innovative games*, Morgan Kaufmann.

- Hazlett, R.L., 2006. Measuring emotional valence during interactive experiences: boys at video game play. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*. pp. 1023–1026.
- Hazlett, R., 2008. Using Biometric Measurement to Help Develop Emotionally Compelling Games. In *Game Usability: Advancing the Player Experience*. Morgan Kaufmann, pp. 187-205.
- Huck, S.W. & Cormier, W.H., 1996. *Reading statistics and research*, New York: HarperCollins.
- Ivory, J.D., 2006. Still a Man's Game: Gender Representation in Online Reviews of Video Games. *MASS COMMUNICATION & SOCIETY*, 9(1), pp.103–114.
- Jenkins, R., Lee, M., Archambault, R., Hebard-Massey, S., Short, G., 2010. *The Influence of Professional Critic Reviews. EEDAR/SMU Behavioral Study*, Southern Methodist University, Guildhall. Electronic Entertainment Design and Research.
- Joeckel, S., 2007. The Impact of Experience The Influences of User and Online Review Ratings on the Performance of Video Games in the US Market. In *DiGRA*. pp. 629–638.
- Kahneman, D. & Tversky, A., 1972. Subjective probability: A judgment of representativeness. *Cognitive psychology*, 3(3), pp.430–454.
- Koeffel, C. Hochleiter, W., Leitner, J., Haller, M., Geyen, A., Tscheligi, M., 2010. Using Heuristics to Evaluate the Overall User Experience of Video Games and Advanced Interaction Games. In *Evaluating User Experience in Games*. Springer, p. 236–260.
- Korhonen, H. & Koivisto, E.M.I., 2006. Playability Heuristics for Mobile Games. In *MobileHCI*. pp. 9-16.
- Korhonen, H. & Koivisto, E.M.I., 2007. Playability heuristics for mobile multi-player games. In *Proceedings of the 2nd international conference on Digital interactive media in entertainment and arts*. p. 35.
- Korhonen, H., Paavilainen, J. & Saarenpää, H., 2009. Expert review method in game evaluations: comparison of two playability heuristic sets. *Proc. of MindTrek 2009*, pp.74–81.
- Lang, P., 1980. Behavioral treatment and bio-behavioral assessment: Computer applications. In *Technology in Mental Health Care Delivery Systems*. Norwood, NJ: Ablex Pub. Corp., pp. 119-137.
- Larsen, J.M., 2008. Evaluating User Experience—how game reviewers do it. In *Evaluating User Experiences in Games, Workshop at CHI*.
- Livingston, I.J., Mandryk, R.L. & Stanley, K.G., 2010. Critic-Proofing: How Using Critic Reviews and Game Genres can Refine Heuristic Evaluations. In *Future Play 2010*. Vancouver, Canada, pp. 72-79.

- Livingston, I.J., Nacke, L.E. & Mandryk, R.L., 2010. Critic-Proofing: Robust Validation Through Data-Mining. In *Fun and Games 2010 Workshop*. Playability and Player Experience. Leuven, Belgium: NHTV Expertise Series, pp. 81-94.
- Livingston, I.J., Nacke, L.E. & Mandryk, R.L., 2011a. Influencing Experience: The Effects of Reading Game Reviews on Player Experience. In *International Conference on Entertainment Computing 2011*. Vancouver, Canada.
- Livingston, I.J., Nacke, L.E. & Mandryk, R.L., 2011b. The Impact of Negative Game Reviews and User Comments on Player Experience. In *Sandbox '11: The 6th ACM SIGGRAPH*. Vancouver, Canada.
- Mandryk, R.L., 2008. Physiological Measures for Game Evaluation. In *Game Usability: Advice from the Experts for Advancing the Player Experience*. Morgan Kaufmann.
- Mandryk, R.L., Atkins, M.S. & Inkpen, K.M., 2006. A continuous and objective evaluation of emotional experience with interactive play environments. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*. Montréal, Québec, Canada: ACM, pp. 1027-1036. Available at: <http://portal.acm.org/citation.cfm?id=1124772.1124926> [Accessed July 29, 2010].
- Mandryk, R.L., Inkpen, K.M. & Calvert, T.W., 2006. Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour & Information Technology*, 25(2), p.141.
- Nacke, L.E., 2009. From playability to a hierarchical game usability model. In *Proceedings of the 2009 Conference on Future Play on@ GDC Canada*. p. 11–12.
- Nacke, L. E & Lindley, C.A., 2010. Affective Ludology, Flow and Immersion in a First-Person Shooter: Measurement of Player Experience.
- Nielsen, J., 1992. Reliability of severity estimates for usability problems found by heuristic evaluation. In *Posters and short talks of the 1992 SIGCHI conference on Human factors in computing systems*. p. 130.
- Nielsen, J., 1995. Severity ratings for usability problems. *Papers and Essays*.
- Nielsen, J., 2005. 10 Heuristics for User Interface Design. Available at: http://www.useit.com/papers/heuristic/heuristic_list.html [Accessed July 29, 2010].
- Nielsen, J., 2007. Heuristic Evaluation: How-To. *Use It*. Available at: http://www.useit.com/papers/heuristic/heuristic_evaluation.html [Accessed August 3, 2010].
- Nielsen, Jakob & Mack, R.L., 1994. *Usability Inspection Methods* 1st ed., Wiley.

- Nutt, C., 2010. Back In Space: BioWare On Mass Effect 2. *Gamasutra*. Available at: http://www.gamasutra.com/view/feature/4251/back_in_space_bio%20ware_on_mass_php,%202010. [Accessed July 29, 2010].
- Pennebaker, J.W., Chung, C.K., Ireland, M., Gonzales, A., Booth R.J., 2011. LIWC. Available at: <http://www.liwc.net/> [Accessed July 29, 2010].
- Pinelle, D., Wong, N. & Stach, T., 2008a. Heuristic evaluation for games: usability principles for video game design. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*. pp. 1453–1462.
- Pinelle, D., Wong, N. & Stach, T., 2008b. Using genres to customize usability evaluations of video games. In *Proceedings of the 2008 Conference on Future Play: Research, Play, Share*. pp. 129–136.
- Pinelle, D., Wong, N., Stach, T., Gutwin, C., 2009. Usability heuristics for networked multiplayer games. In *Proceedings of the ACM 2009 international conference on Supporting group work*. p. 169–178.
- Pratto, F. & John, O.P., 2005. Automatic Vigilance: The Attention-Grabbing Power of negative Social Information. *Social cognition: key readings*, p.250.
- Ravaja, N., Saari, T., Salminen, M., Laarni, J., Kallinen, K., 2006. Phasic emotional reactions to video game events: A psychophysiological investigation. *Media Psychology*, 8(4), p.343–367.
- Reinstein, D.A. & Snyder, C.M., 2005. The Influence of Expert Reviews on Consumer Demand for Experience Goods: A Case Study of Movie Critics. *The journal of industrial economics*, 53(1), pp.27–51.
- Russell, J.A., 1980. A circumplex model of affect. *Journal of personality and social psychology*, 39(6), p.1161.
- Sundar, S.S., 2007. The MAIN model: A heuristic approach to understanding technology effects on credibility. *The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning*, pp.73–100.
- Sundar, S.S., Knobloch-Westerwick, S. & Hastall, M.R., 2007. News cues: Information scent and cognitive heuristics. *Journal of the American Society for Information Science and Technology*, 58(3), pp.366–378.
- Sundar, S.S., Oeldorf-Hirsch, A. & Xu, Q., 2008. The bandwagon effect of collaborative filtering technology. In *CHI'08 extended abstracts on Human factors in computing systems*. pp. 3453–3458.
- Sundar, S.S., Xu, Q. & Oeldorf-Hirsch, A., 2009. Authority vs. peer: how interface cues influence users. In *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*. pp. 4231–4236.

- Tassinari, L.G. & Cacioppo, J.T., 2000. The skeletomotor system: Surface electromyography. *Handbook of psychophysiology*, 2, pp.163–199.
- Tversky, A.K. & Kahneman, B., 1974. Judgment under uncertainty: Heuristics and biases. *Science*, 185(4754), pp.1124–1131.
- Yi, Y., 1990. Cognitive and affective priming effects of the context for print advertisements. *Journal of Advertising*, 19(2), pp.40–48.
- Zagal, J.P., Ladd, A. & Johnson, T., 2009. Characterizing and understanding game reviews. In *Proceedings of the 4th International Conference on Foundations of Digital Games*. Orlando, Florida: ACM, pp. 215-222. Available at: <http://portal.acm.org/citation.cfm?id=1536513.1536553> [Accessed August 5, 2010].

APPENDIX A

AFFECTIVE TEXT

A.1. Negative Game Reviews: The Fancy Pants Adventure

A.1.1. Review 1

The Fancy Pants Adventure (FPA) is an online platform adventure game. You take on the role of a boring and crappy animated character that, as one might have guessed, wears fancy pants.

The game takes the most adverse pages from the book of the now defunct Sonic the Hedgehog, namely speed and irritating gameplay. The main character lurches through lackluster levels littered with the occasional obstacle or enemies. Sometimes there is unremarkable garbage thrown carelessly about the godawful levels like the unsatisfying springboard, and some hapless attempts at half-pipes. The game has a fast pace and is burdensome to control – it slows down to let you explore the world, but that’s a humdrum experience. There are several ugly areas hidden obnoxiously, and the effort you’ll waste finding them an utter waste. Did you see the developer’s name inscribed in the stones? Me either, so why bother?

As you explore the worlds in this game, you'll come across a few unimpressive items, namely floating swirls and trophies. Swirls unoriginally refill your health and give you an extra life for every 100 you waste your time collecting. There are cheesy trophies to be collected and viewed from the main room of the game. It’s all very blasé.

Visually and audibly the FPA is sad compared with many similar platform games. The music and sound effects are bad, only serving to distract from the already annoying experience. The art is simple and it shows. The animations are run of the mill, and add little to the game.

Nothing comes together in this faulty game, making for a game that's irksome and way troublesomely easy.

FPA is bad. I really frustrated with this game; it's boring, the controls are too finicky, and it's ugly. There are a lot of flash games out there; most of them are bad. The uninspired FPA fall into that group. Skip this horrendous platformer.

A.1.2. Review 2

Fancy Pants Adventure (FPA) is a completely disillusioned attempt at an unoriginal platforming game. The worlds are mundane and uninspired, and movement and control are nearly useless. It's all terribly agitating that this type of game could be executed so poorly.

The object of the game is to make your way through some dreadful levels and beat the cheesy and dolefully implemented end boss. Along the way you'll face a few lousy enemies (spiders and rats), collect little swirls (equivalent of coins), and breeze through some abhorrent and repetitive level designs. You get no real sense of speed when moving around, and it feels like an alarming bad sonic game. The surfaces in the game have plenty of awkward curves, meaning you will run up the sides of walls, roll down hills, and find the levels unnecessarily hamper your movement. I spent 10 miserable minutes trying to land on a surface that I just kept missing due to the frustrating physics. The character himself is an uninteresting stick figure with crappy animation in an attempt to be current. There are a lot of little troubling things that make this game irritating. Uncontrollable and frustrating jumping mechanics, bad hit boxes, and the controls are dreadfully awkward. The whole experience is unpleasant and even the extras such as trophies don't make this shabby platformer worth replaying. It's the combination of poorly conceived physics, pathetic animations, and the heartbreaking level design that make FPA not worth playing.

Graphics are lousy and unstylish. It looks like some kid with a crayon tried really hard to draw, but ended up just making a mess. The game also has an unimpressive soundtrack to go along with it, but it really doesn't fit the atmosphere of the game, and just becomes distracting. The collection of trophies adds little to the experience, and I didn't have any need to actually look for them since they were so poorly hidden.

FPA is a flash game, and a crappy one at that. It's boring and annoying at the same time; I don't know what the developers were thinking when they designed this game. There are so many other indie flash games out there, why not spend time with one of them?

A.1.3. Review 3

Fancy Pants Adventures is unfortunately not worth playing. Even the die-hard platform fans should avoid this embarrassment.

FPA is an unpleasant experience due primarily to its inconsistent and lackluster platforming gameplay. Gamers guide Fancy Pants Man through a series of disappointing and unimpressive stages that feature half-loops, cliffs, hills, and other such uninteresting environmental features, usually placed badly and inconstantly throughout each level. There are lame little collectible "coins" in the form of squiggles and hidden trophies strewn across the entire game, but you don't need to look too hard to find these unimaginative doodads (sorry hardcore completists).

What makes FPA so boring is its unpleasant platforming. There are a couple of uninspired giant spiders and other enemies that gamers can take down, but for the most part, FPA is about jumping from ledges, bouncing off of springs, and running through mediocre levels. It's a simple formula, so it's unfortunate that FPA did such a terrible job at making the embarrassing gameplay even remotely appealing.

FPA has a tired hand-drawn look to it. The color palette consists of depressingly dull pastel hues, and there are scribbles here and there that fail to add anything to the look the game seems to be touting. Fancy Pants Man himself looks like a hand-drawn stick figure, and this fails to help him stand out from his gloomy environment.

Throughout each of the painfully poor levels in FPA, gamers will hear an irritating, ambient soundtrack that just distracts from the gameplay and the overall experience. The music in the game has that unstylish, trying-too-hard indie hipster vibe, and it only serves to degrade the game's already troubling vibe even more.

Not only is FPA a boring complete nonsense platformer, it relies on a dated and annoying gaming idea that should be left to expire in piece. It's so frustrating that it would be hard to urge even the most die-hard platform gamers to play the game. In many cases, simplicity can be harmful to a video game, and in the case of FPA even the simplicity is done at distressingly poor level.

A.1. Negative Game Reviews: Tag: The Power of Paint

A.1.4. Review 1

Tag: The power of paint is an irritating puzzle anthology intricately played from first-person perspective with the assistance of a foul gimmick too bad to warrant attention. The game is bad. Everything about this pathetic game is unpleasant, from the boring gameplay to the uninspired grungy art. The hapless super-soaker paint gun squirts paint; three different hideous colors of paint each granting the player a cheesy and boring ability. Green paint makes you jump, red paint lets you run fast, and the sticky blue paint lets you walk up walls tediously, and across mouldy ceilings, a mechanic that will leave you with a sickly sense of vertigo. Left-clicking

produces a distressing stream of the currently selected color, coating whatever happens to be ahead of you. Right-clicking provides a stream of water, which washes away any paint it touches. You will spend half your time wearisomely trying, unsuccessfully, to get the paint to go where you want it too. Difficulty is always important in any puzzle game, and Tag is no exception. The puzzle difficulty level in Tag is way off base with everything being annoying easy, and that makes for dreary game play. The more intricate the bothersome puzzle, the more annoying the level becomes the pathetic game is never hard; the ill-designed paint mechanic is just frustrating to use. Part of Tag's defectiveness lies in the dubious cheese-crumbs trail it leaves; everything is just too easy. With the help of ridiculous hints blatantly spread through each level, the intermittent game never takes off but remains tedious and lackluster throughout. Tag is a generic, irritating, and generally unpleasant rip-off of other games.

The world of Tag is an unattractive, cell-shaded world of uninviting grays and whites. Your only method of tedious progression is the tiresome painting mechanics, and your often-insignificant puzzle solving ability. Puzzle solving tries to be an end in itself – but fails, and the clunky game's wide-open levels are ugly and uninspired. If there is an ill-favored attempt to produce quality it never appears, leaving an abysmal and hideous experience that has been done better in a hundred other games.

The overall style of Tag makes the game unsightly and tragic in its flawed execution. There is a lack of consistency, there's the grungy graffiti-like text and a depressing gray urban setting, which really does nothing to help the crappy gameplay experience. The simplicity here is ineffective-while it works in other games there is literally nothing in this foul game worth doing. The awful sound effects used in Tag are really substandard as well, and are easily forgotten,

often becoming annoying. It all adds up to an experience that's too boring and too easy to be described as anything other than mediocre.

Tag: the power of paint is an unimpressive failure. It fails to prove the creative viability of games that are low-rent/casual/arcade all-or-nothing wager in the monolithic triple-A market. Above all, I was annoyed with this game. Do not play it, it's really awful.

A.1.5. Review 2

The object of each level of Tag is to get from where you've started to the mystical blob at the end of the level, an incredibly boring task. Sometimes the buggy paint will be there for you to pick up, other ill-designed levels force you to figure out how to tediously get to the paint cans before you can actually play this foul game, which is a lame endeavor just by itself. It almost hurts me physically to talk about this stupid game; it is that bad. I was annoyed almost the entire time I spent with struggling with this demented game.

Each of the three colors of ugly paint causes you to perform some uninspired action when you come into contact with the paint. Green makes you jump, red makes you move fast, and blue causes you to stick to the paint, which will enable you to climb walls. The most tiresome was the green paint – it basically gives you a pathetic ability that every other game gives you by default, I mean really who can't jump? I became worried by the ridiculous lack of jump when I started playing, and my concern was proven regrettably warranted.

The shoddy game relies on simple unsound mechanics, poorly arranged, which causes an overall loathsome experience. What poorly exists is a dreary experiment in physics and puzzle solving. Since there are a limited number of odious options here, you have to decide whether to jump, run really fast, or climb walls. There are pre-painted surfaces making you weakly think through your next move; if you don't you're going to fall miserably to your death, killing you

and frustratingly sending you back to the last checkpoint – did I mention there are no saves? The atrocious challenge level in Tag is infuriating compared with other games. The game gives little or no hints to help guide the desperate player; however, given that the bad game is already too easy, the hints are an unnecessary little touch.

There are buggy checkpoints, rather than saves, that eventually allow you to keep progressing instead of having to start to at the beginning every time you misstep – personally I hate checkpoints.

Stylewise the game looks unpalatably ugly. The world is poorly designed and scantily colored in a way that makes Tag unpleasant – say it with me folks: grey-scale looks bloody nasty! There has been a lot of wasted potential on the part of the pathetic designers resulting in a cheerless style – I mean who ever thought a boring gray world would look anything but bloody ugly? The art and music really work badly together and just add to the dull game play experience.

I really hate this game. It looks ugly, it's boring, and the puzzles are too easy. It's simple in so many ways, it uses basic and boring mechanics to craft an uninspired gameplay experience for the player. You need to avoid this unimpressive game. It is really bad.

A.1.6. Review 3

Tag: the power of paint gives players a generic boring puzzle gameplay experience. Set in an ugly cell-shaded, grayscale, urban landscape, Tag gives players a tiresome way to intricately move through the unremarkable levels – levels that are so cumbersome and elementary they fail to challenge the players' spatial awareness even simplistically. Tag is a loathsome and monotonous game that packs some tediousness into a small package.

Tag uses a stupid concept that's very simple to wrap your head around without actually seeing it in action – walking on paint gives you shoddy powers. The dismal first few levels of Tag introduce you to the unimpressive applications your paint provides, I was really annoyed with how easy the game was. Certain surfaces are not paint-able (like chain-linked fences) – an atrocious feature that does not make sense – painfully forcing you to use the generic ability that each paint offers. Other ill-designed elements, such as moving platforms, are strangely incorporated into the moronic levels in increasingly lame ways. The numskulled puzzles are always too simple; there are always little pathetic hints that are not required, showing the player where they need to head next.

The sound design, while minimal, is bloody horrible. My six-year old could have done a better job. There's also something generic and dull about the sound of the paint gun. You are not shooting at anyone or necessarily shooting a gun that requires a louder sound, but this particular gun sound is terribly unpleasant and almost annoying in its quality – by the end of this disaster I had grown to hate it. Both the sound and the art direction are terrible and do nothing to the level of consistency in the game.

Tag is a pathetic game that should only be played by anyone interested in generic, and brainless gameplay. It's gray and everything is done poorly. I would never recommend Tag: The Power of Paint to anyone unless I hate him or her. It's ugly, boring, and too simple, a truly unpleasant experience.

A.2. Positive Game Reviews: The Fancy Pants Adventure

A.1.7. Review 1

The Fancy Pants Adventures (FPA) is an online platform adventure game. You play a hip and uniquely animated character that, you guessed it, wears fancy pants.

The game takes the most enjoyable pages from the book of Sonic the Hedgehog, namely speed and out of reach secrets. The main character dashes through brilliant open levels littered with the occasional obstacle or enemies. There are fun things like springboards to send you skyward, awesome half-pipes to run around, and hills that propel you in many directions. The game has a fairly fast pace and is challenging but slows down to let you explore the unique world. In fact, there are several fun hidden areas you can find if you take the time to look around. Did you see developer's name inscribed in the stones?

As you explore the sleek worlds in this game, you'll come across tons of items, namely floating swirls and trophies. Swirls refill your health and give you an extra life (measured in Fancy Pants, of course) for every 100. Trophies can be collected and viewed from the main room of the game.

Visually and audibly the FPA is excellent. The music and sound effects add a lot to the overall fun experience. The art is simple but well animated adding a clever whimsy to the game. Everything comes together making for an enjoyable and challenging game. Every game has room for improvement, however FPA doesn't leave much room at all. All we can do is hope for more, because when I finished the game I was left wanting more.

FPA is a complete and polished experience. I really love this game; it's enjoyable, it's challenging, and it's brilliant. There are a lot of flash games out there, most of them are bad, but Fancy Pants doesn't fall into that group. Are you ready to try on your fancy pants?

A.1.8. Review 2

Fancy Pants Adventure (FPA) is an absolutely brilliant platforming game. The worlds feel open and free, and movement and control are nearly flawless. It's all great fun.

The object of the game is simply to make your way through some levels and beat the end boss-not a complex premise, but one that has stood the test of time. Along the way you'll face a few unique enemies, collect little swirls (equivalent of coins), and work your way through the varied level designs. You get a real sense of speed when running around, but not so much that it feels like an out-of-control sonic game. The surfaces in the game have plenty of delightful curves, so this means you can run up the sides of walls and roll down hills. The character himself is a stick figure but is animated very fluidly, and there are a lot of little touches that make this game enjoyable. There's a great little turning animation when you switch from forward to backward, and if you jump in the middle of the animation you'll do an awesome backflip. When coming out of a loop your character will continue moving in the angle and momentum they had before, making curve-to-curve jumps look neat. It's the combination of fun physics, awesome animations, and the enjoyable level design that make it entertaining to just run and jump around in a level without trying to go anywhere. FPA is able to pack a lot on gameplay into a relatively small package. At no point did I feel bored while playing. There was always something to do, or somewhere to go. Even after playing through twice I still feel that I have missed something – a feeling that makes me want to play again.

Graphics are simple but stylish. The game also has a great soundtrack to go along with it that really fits the casual fun atmosphere of the game. The game also features a good collection of trophies to find and collect, adding replay value as most of them are pretty challenging to come across. We all know that these additions are simple, but without them the game would feel unfinished.

FPA is an absolute must-play flash adventure game. You'll likely find yourself playing over and over again. It's fun and unique, and shows that speed platformers can still be awesome.

A.1.9. Review 3

Fancy Pants Adventures is extremely fun and worth playing by platform game fans.

FPA is easy to get into due primarily to its unique platforming gameplay. Gamers guide Fancy Pants Man through a series of stages that feature half-loops, cliffs, hills, and other challenging environments. There are also little collectible “coins” in the form of squiggles and hidden trophies strewn across the entire game, so a little exploration is definitely necessary for hardcore completists.

What makes FPA so much fun is its pure, unadulterated and extraordinary platforming. There are giant spiders and other enemies that gamers can take down, but for the most part, FPA is about jumping from ledges, bouncing off springs, and running through levels at high speeds. It’s a simple formula, but an extremely fun one at that.

FPA has a wonderful hand-drawn look to it. The color palette consists of great pastel hues, and there are scribbles here and there that add a cool touch to the game’s artistic look. Fancy Pants Man himself looks like a hand-drawn stick figure, fitting with the game’s look, and giving him charm and personality unique to his character. The game is simple and packs a decent amount of gameplay. It’s hard to find fault with such an elegant design, and any faults found are quickly forgotten due to the games charming nature.

Throughout each of the levels in FPA gamers will hear a cool, ambient acoustic guitar style soundtrack that can instantly be associated with the game’s unique art style. The music in the game has that stylish and of-the-moment indie hipster vibe, and it emphasizes the game’s indie feel even more. If only more platformers could be this charming.

Not only is FPA a fun no-nonsense platformer, it is a cool and unique gaming experience that no one should miss. It is hard not to urge gamers to play the game over and over. In many

cases, simplicity can be harmful to a video game. In the case of FPA, however, simplicity is one of the keys to the gameplay experience.

A.3. Positive Game Reviews: Tag: The Power of Paint

A.2. Review 1

Tag: The power of paint is an amazing puzzle anthology played from first-person perspective with the assistance of a gimmick too good to neglect. You are given a delightful super-soaker styled paint gun squirts paint; three different colors of paint each granting the player a fun and unique ability. Green paint makes you jump, red paint lets you run fast, and the sticky blue paint lets you walk up walls, and across ceilings, which means that every corner in sight is also brilliantly within reach. Left-clicking produces a colorful stream of the currently selected paint coating whatever happens to be a head of you. Right-clicking provides a stream of water, which conveniently washes away any paint it touches. Difficulty is always important in any puzzle game, and Tag does great work balancing fun and challenge. The puzzle difficulty in each level is just right and makes for enjoyable game play. The more intricate the puzzle, the more fun the level becomes—part of Tag's greatness lies in the cheese-crumbs trail it leaves. With the help of hints cleverly spread through each level, the game becomes quickly addictive and remains marvelously satisfying. Tag is truly a unique, delightful, and fun game.

Tag is set in a wonderful cell-shaded world of grays and whites. Your awesome painting skills are the key to your progression and puzzle solving. The latter is an end in itself, and the game's wide-open levels are enjoyable to crack again and again.

The overall style of Tag makes the game very fun. I really enjoyed the level of consistency, from the graffiti-like text to the gray urban setting; It really helps to bring the great gameplay experience together. There is a simplicity here that is delightful, and the things the

game does it does exceptionally well. The sound effects used in Tag are high quality as well, and really add something awesome to the game. It all adds up to a very pleasant experience.

Tag: the power of paint is an awesome game, and proves the creative viability of games that are neither low-rent/casual/arcade all-or-nothing wager in the monolithic triple-A market. Above all, know that I love this game, cannot wait for more, and recommend this game very highly, it is really just that good.

A.2.1. Review 2

The objective of each amazing level of Tag is to get from where you've started to the mystical blob at the end of the level, an incredibly fun and challenging task. In order to do this you'll be using the W-A-S-D or arrow keys for movement and the mouse for looking around, as well as a gun that sprays paint. Sometimes the paint will be right there for you to pick up; other levels challenge you to figure out how to get to the paint cans before you can use them, which can be a fun puzzle just by itself.

Each of the three unique colors of paint allows you to do something unique when you come into contact with the paint. Green makes you jump, red makes you move fast, and blue causes you to stick to the paint, which will enable you to climb walls. I had the most fun playing with the blue paint, really taking the time to explore the open levels – something I would recommend as a really fun exercise.

The game relies on simple mechanics, expertly arranged to provide a great experience. What exists is a fun experiment in physics and puzzle solving. Since there are a limited number of options here you have to decide whether to jump, run really fast or some combination of these great abilities. The pre-painted surfaces require you to really think through your next move, providing a wonderful new level of dimensionality to the game. For example, do you really want

to land on a green surface after a jump? The challenge level in Tag is excellent. The game offers little visual hints to help guide the player without revealing how to solve the puzzle – a wonderful little touch. The gameplay is very stimulating challenging you spatially in ingenious ways; the blue paint especially will have you tilting your head left and right as you find your bearings.

There are cool little checkpoints that allow you to keep progressing instead of having to start all over every time you misstep. The checkpoints also provide the player with great guidance. Seeing a checkpoint on top of a building, you will say to yourself, “Maybe I should be trying to get up there.”

Stylewise, the game looks amazing. The world is designed and colored in a way that makes Tag really enjoyable. There has been a lot of great effort on the part of the designers to provide a wonderful artistic style. The art and music really work well together, providing a brilliant gameplay experience.

I really love this game. It looks fantastic and has challenging puzzles. While simple in many ways, it uses its unique mechanics to craft an enjoyable gameplay experience for the player. You need to play this awesome game. What are you still reading for? Go get it now!

A.2.2. Review 3

Tag: the power of paint gives players a uniquely puzzling gameplay experience. Set in a pleasantly cell-shaded, grayscale, urban landscape, Tag introduces players to creative new ways of moving through the awesome levels that challenge the players’ perspective and spatial awareness. Tag remains a unique and fun game that packs a huge amount of delight into a small package.

Tag uses a concept that's challenging to wrap your head around without actually seeing it in action. The instantly addictive first few levels of Tag introduce you to the innovative applications your paint provides, eventually giving you all 3 colors of paint so you can decorate anywhere you want. Well, not quite anywhere. Certain surfaces are not paint-able (like chain-linked fences), and this is one way that subsequent levels become more challenging, allowing you to use the unique ability that each paint offers. Other puzzle elements, such as moving platforms, are incorporated into the levels in increasingly fun ways, but the real complexity and genius of Tag lie in combining paints to solve the puzzles. However, the puzzles are never too hard; there are always little hints that are great at showing the player where they need to head next.

The sound design, while minimal, is great. There's also something uniquely and tremendously satisfying about the sound of the paint gun. You are not shooting at anyone or necessarily shooting a gun that requires a powerful sound, but this particular gun sound is appealing and almost whimsical in its sonic quality. Both the sound and the art style add a first-rate level of consistency to the whole game – it is clever how these little details complete the game.

Tag is a wonderful game that should be played by anyone interested in unique, enjoyable gameplay. I would recommend Tag: The Power of Paint to everyone. It's unique, fun, and challenging, a truly wonderful experience.

A.4. Negative Movie Reviews: Vulgar

A.2.3. Review 1

Starting out as a witless, unimpressive rip of Bobcat Goldthwait's "Shakes the Clown," then veering toward what appears to be a disappointing rape-revenge drama, writer-director

Bryan Johnson's "Vulgar" is the type of seriously awful movie that sounds so bad, it might not be. Sitting through this miserable film is another matter, though. Final credit thanks exec producer Kevin Smith, "without whom I'd still be working at the car wash and you wouldn't be reading this." Need more be said about the grotty quality?

Brian O'Halloran disgraces himself in a dull role as William aka Flappy, a professional clown, Will Carlson. His birthday-party business is a bust, his cruel life is a "dismal failure." Desperate for cash, Will decides to sideline as a pathetic cross-dressing gag act for bachelor parties. A truculent decision. His first job gets him beaten, gang-raped and otherwise humiliated by a middle-aged perv and his two inbred sons. Later, miserable Will saves a child hostage, earning his own national TV kid show -- which in turn draws the attention of the sanguinary goon trio, whose scrubby blackmailing scheme leads to protag's bloody vengeance.

That's the long and short of "Vulgar," an ugly, revolting movie that marks the writing-directing debut disaster of Bryan Johnson, an unimportant actor who has appeared in several Kevin Smith movies ("Mallrats," "Dogma"). The misguided Smith executive-produced this miserable and plays a poor role as a TV producer.

It's hard to know if Johnson intended this crappy, paltry material as darkly funny, since there's nothing to laugh at. The dreadful rapist has two sons, one short and lisping (Matt Maher), the other enormous and cretinous (Ethan Suplee), who assist him in gross acts of sexual terrorism. He also has a young wife (Susanna Jolly) and daughter (Ceton Tate) who despise him.

Johnson, who also plays Will's best friend, has a twisted, dingy mind and no evidence of talent whatsoever. It's impossible to imagine why Lions Gate, the indie distributor that released "Monster's Ball," would bother with this inferior, terrible garbage.

Inexplicably mixing lamer-than-lame "bad taste" comedy with much worse traumatized-assault-victim histrionics, pic's only value lies in viewer weighing whether pic is primarily a.) offensive b.) amateurish c.) pathetic or d.) a cry for help. A terrible and dull movie that leaves a bad taste in your mouth. Do not watch this shit!

A.2.4. Review 2

With the flood of marginal independent films that have hit New York in the last few years, it has become more common to see critics staggering out of the screening room, shaking their heads in astonished disbelief. Bryan Johnson's "Vulgar," though, provokes a whole new level of incredulity.

"Vulgar," which Mr. Johnson wrote, directed and edited, is the relentlessly unpleasant tale of a New Jersey clown, Will Carlson (Brian Christopher O'Halloran), who thinks only of bringing happiness to children — a thought that's hard to maintain in the run-down, drug-infested neighbourhood where he lives.

Birthday parties being in short supply, Will decides to augment his income by working the bachelor-party circuit. He invents a new character, a stubble-faced, transvestite clown named Vulgar, and places an ad in a newspaper. His first gig leads him to the most sordid motel west of the Hudson, where he is gang-raped by three dingy mental deficients who seem to have migrated north from "The Texas Chainsaw Massacre": a sadistic father (Jerry Lewkowitz) and his two pathetically servile sons (Ethan Suplee and Matt Maher).

Will has touched bottom, as he confesses to his friend, Syd (played by Mr. Johnson with more assurance than he shows as a filmmaker). But when he rescues a young girl from her abusive, gun-wielding father, he becomes a national celebrity and is offered a network children's show of his own.

Then the rapists resurface, with a complete videotape of the assault. Knowing his career will be ruined unless he recovers the tape, Will arms himself with a couple of untraceable guns and goes out in search of his tormentors.

Too campy to work as straight drama and too violent and sordid to function as comedy, "Vulgar" is, truly a one-of-a-kind work. Mr. Johnson went to high school with the director Kevin Smith, who was godfather to this film and contributes a cameo as a gay television producer. As a movie, "Vulgar," which opens today in Manhattan, San Francisco and Los Angeles, is quite a tribute to friendship. It certainly has no other apparent reason to exist.

A.2.5. Review 3

Imagine Kevin Smith, the blasphemous bad boy of suburban Jersey, if he were stripped of most of his budget and all of his sense of humor. The result might look like Vulgar, a chintzy "transgressive" feature that dares you to find it tasteless and ugly and sordid and a lot of other things that have long since become commonplace in polite society. Financed by Smith's View Askew Productions, and written, directed, and coedited by his protégé (read: old teenage chum) Bryan Johnson, the movie is built around a spectacularly unpleasant sequence in which the title character, a loser who works as a party clown, hires himself out to a bachelor party, where he is gang-raped by a trio of scummy lowlifes who look like they stepped out of some long-lost John Waters version of "Deliverance."

And that's the sad part of the movie. Brian O'Halloran, one of the stars of "Clerks," attempts to act up a storm as the beleaguered victim, who in a fearless satirical twist soon emerges as...a media star! "Vulgar," a movie that seems to have been designed to shock the kind of scabby people who once attended grungy midnight shows. The only thing shocking about it,

however, is the degree to which self-congratulatory gutter exhibitionism has become the degraded ash end of indie "edge."

A.5. Negative Movie Reviews: The Singing Forest

A.2.6. Review 1

There isn't much singing in "The Singing Forest," but there is sure a lot of unpleasant whining and much other stuff to hate. Writer-director Jorge Ameer's pathetic bare-bones vanity project has one incredible premise, one that is so ridiculous and over-the-top, it would have been difficult to take his scabby film seriously even if the rest of the film's elements were top-notch (they are not).

Christopher (Jon Sherrin), a battered magazine journalist who is drinking heavily after the death of his wife of 22 years, goes to his grungy daughter's house to meet her seedy fiancé, Ben (Craig Pinkston). Christopher has become convinced, after a recent visit to a psychic, that he was a German resistance fighter hanged by the Nazis during World War II. And when he meets Ben, he is equally convinced that his sorry daughter's poor fiancé is the reincarnation of his gay lover. At this point, you might want to consider whether it is actually worth suffering through the rest of this inferior, terrible movie, no matter what your motivations were to see it.

So the creepy, boozing and much older Christopher sets about meanly seducing his daughter's dim-witted lover. If you do not dislike the terribly executed script yet, wait until the movie becomes painstakingly worse to watch by the minute.

Ameer is obviously riffing off a tormenting scene in the documentary about the persecution of homosexuals by the Nazis, "Paragraph 175." The "singing forests" described in that stirring documentary were the sordid screams of horrified victims mounted on poles. A

bestly picture reprehensible, shameful deeds done in history's darkest hour appallingly burnt into your distraught memory.

If that seems impurely tasteless, it is. So are the various other sundry and revolting plot revelations -- for example, it turns out that Christopher met his grungy wife and conceived Destiny, his daughter, by raping her nastily. What a vile guy. Are you sick of this rubbish, yet?

Most perplexing and pathetic of all is the perverted portrayal of Destiny (Erin Leigh Price). Here is a young woman whose mother has recently died, and whose dopey fiancé is sleeping with her sleazy father - and she seems OK with it. No emotional scenes, no arguments with Dad, no commiserating over her mother's painful passing. What an ugly and vicious take on a potentially interesting role. Apparently, she's just in it for her name.

A.2.7. Review 2

I hate "The Singing Forest." Jorge Ameer's degenerate, abominable drama about reincarnation gone wrong, is unlikely to win any daffy converts. Awkwardly staged and squalidly edited and fitted out with an overly intrusive, abhorrent score drawn primarily from classical music, this rubbish movie consistently subverts the earnest efforts of its pathetic cast.

A middle-aged, widowed writer, Christopher (Jon Sherrin), has long been haunted, even before the birth of his dim-witted daughter, Destiny (Erin Leigh Price), by the miserable belief that in an earlier life he was a dumb gay German who with his lover became victims of the horrible Holocaust. Destiny, who he has not seen since her silly college days, has now informed him of her upcoming marriage. Creepy revelations soon abound.

Father and daughter are reunited but, no sooner does Christopher meet his soon-to-be son-in-law, foolish Ben (Craig Pinkston), than he senses that Ben is the cockeyed reincarnation

of the gay German's lover. If this sounds like a terrible stretch, it's only the beginning of a detestable, imbecile series of disgusting events.

If nothing else can be said of "The Singing Forest," it is assuredly fearless in oafishly defying credibility at every turn and on every atrocious level. I am not sure anyone else would want to suffer through this inferior, god-awful mess of a movie. There is no single obscene reason on earth to watch crap like this.

A.2.8. Review 3

Some shitty films are best encapsulated via imaginary tabloid headlines, which in "The Singing Forest's" case would be: "Daughter's Fiancee Is My Reincarnated Gay Nazi Lover, Dad Claims." Given a scandalous conceit that jaw-dropping, it is sad to report writer-director-producer Jorge Ameer's heinous feature mistake provides scant value, intentional or otherwise. Talk-ridden, exceedingly awkward drama is a tough sit even at just 72 distressing minutes, with a shakier grasp on the rudimentaries of storytelling and more lousy tech work than evident micro-budget can plaintively excuse. Filmmaker is self-distributing; prospects are marginal at best. Pathetic movies cannot get much worse than this sordid, despicable tale of gay Nazi reincarnation gone wrong.

Depressed, beer-guzzling and inattentive to miserable work since his wife of 22 years died some grim months earlier, battered journalist Christopher (Jon Sherrin) goes to visit dull-witted only child Destiny (Erin Leigh Price) when she's about to marry live-in b.f. Ben (Craig Pinkston). Upon meeting him, however, Christopher realizes grungy Ben, an ugly male model, is the reincarnation of a stupid lover he'd had in a past life (a relationship recently confirmed by a psychic). They were a German-Jewish resistance fighter and "good" Nazi soldier, both horribly executed during WWII.

Not surprisingly, dim-witted Ben is nonplussed by the bad revelation ("I'm not some crazy ghost from the spirit world -- I'm from Des Moines!"), though he also drunkenly tumbles into future-dad-in-law's spare-room bed not once but twice. The kind of complex character writing that might have somewhat ballasted such improbable shenanigans is unpleasantly absent here, making for a deplorable patchwork script with faulty ideas. Ditto any – even generic – atmosphere supportive of love-from-beyond-the-grave fantasy.

Tragically, pic's problems leave no department untouched. Script is awful and shameless in detail and incident. Unbearable dialogue is by turns New Age-y and banally improvised-sounding, the ending simply confusing and unsatisfactory. A key argument is filmed silent, following a derogative scene in which protags' words are semi-drowned out by crashing ocean waves; scenes of people watching television are mysteriously accompanied by no small-screen image or sound. Insulting to anyone with taste in filmmaking really. One can frequently hear wind or fingers thumping against the sound recordist's mike; camera tends to wobble (was no tripod available?); several scenes are soundtracked by calamitous library orchestral excerpts that seem execrable in this faulty context.

Both well-toned male leads are seen fully nude, yet "Forest" bungles even its exploitation card so thoroughly it's unclear whether they actually have sex, derogatory or not. In general, one could say that it is disturbing to watch. The Nazi association (replete with actual disgraceful concentration-camp photos under opening credits) would be offensive if the film weren't so trivial, hapless and dull -- it lacks the juice to render even bad-taste potent.

Despite barely feature length, pic still finds room for gratuitous scene repetitions and long, boring credit crawls.

A.6. Positive Movie Reviews: Good Bye, Lenin

A.2.9. Review 1

"Good Bye, Lenin!" is both a wonderful and touching story of a son's astonishing devotion to his beloved mother and a wry commentary on how not all East Germans were thrilled by the dazzling fall of the Berlin Wall. Set in 1989, just as capitalism trumps communism, the movie comes down equally hard on both systems. Communism may have turned believers into Stepford wives, but at least you weren't bombarded with Coca-Cola signs and otherwise prodded to consume. It's a genial measure of the film's delectable evenhandedness that it's been an amazingly huge hit all over Germany.

The movie centers on the Kerner family, living in drab East Berlin quarters minus Dad, who's gone to the West because of police nuisance over his refusal to be a good soldier. Family matriarch Christiane (The wonderful Katrin Sass) more than merrily compensates for her husband's lapses. She nicely embraces the party line, designing nondescript clothes for paragon women like herself, who would be baffled at the thought of playing up their looks.

The comedy turns decidedly black when Christiane has a heart attack and is in a coma while the wall comes down. When she suddenly awakens eight months later, her son, Alex (The talented Daniel Bruhl), is warned that any shock could cause her heart to give out. His delightful response is to pretend unification never happened. What follows is a wonderfully funny collection of scenes where Alex hauls back the standard-issue furniture his mother remembers and, failing to find her favourite brand of pickles among the array of Western products, combs through garbage cans for old jars so he can switch labels.

The funniest bits evolve from her simple request to watch the news. Initially, Alex is able to comply by playing tapes of pre-unification news shows on a VCR cleverly hidden from view.

But soon Christiane begins noticing inexplicable things, like a red communist banner being replaced by a Coke ad right outside her window. So Alex has to get creative, staging a marvellously funny newscast with an ersatz anchorman cheerfully announcing that Coke turns out to have been a glorious East German invention, a great moment in this excellent movie.

There's a miraculous freedom to brilliant scenes like these that reminded me of Second City in its heyday. Director Wolfgang Becker wrote the astoundingly clever script (along with the talented Bernd Lichtenberg), but I wouldn't be surprised if he allowed his flamboyant actors to lushly improvise on it. At any rate, they sure seem to be having lots of fun, and that just makes for an even greater movie experience.

The leads are particularly fascinating and enjoyable. Sass, who was an endeared big star of East German cinema, brings an appropriate dourness to the precious role. Yet she lets Christiane's maternal side peek through, so you understand why her wonderfully adorable son would instigate such an elaborate ruse on her behalf. Bruhl is a beguiling charmer, able to capture Alex's sweet dilemma as he is caught between two regimes, each with its strengths and weaknesses.

"Good Bye, Lenin!" looks as if it was enthrallingly shot on a shoestring -- sets are minimal, and there aren't a lot of extras hanging around. But even this works fascinatingly to the movie's advantage.

The pitch-perfect title refers to one of those mesmerizing statues the Communists were so fond of constructing to honour their leaders. Adding to Christiane's confusion is the sight of Lenin's head and torso hanging from a helicopter on its way out of town.

A.2.10.Review 2

Loving Grown children protect mom from the delightful political ravages of Burger King and Coca-Cola in an awesome comedy to prove that few, if any, screen premises are impressively outlandish if handled with care. Though a Hollywood sledgehammer approach could really ruin this splendiferous story.

Long-abandoned by her husband for the West, an East German mother (The beautiful Katrin Sass) devotes her paragon life to communism until she suffers a heart attack at a 1989 rally and falls into an eight-month coma. When she awakens, the Berlin Wall has fallen. Her daughter (The talented Maria Simon) is pushing fast food, and her gorgeous son (The handsome Daniel Bruhl) is selling satellite dishes that transmit news broadcasts that'll do the frail recoveree no good.

So it becomes a magnificent madhouse staging and taping phony accounts (for mom's consumption) that delightfully extol East German superiority and rationalize vastly increased traffic over the border. Harder to explain, though, is that eye-catching Coca-Cola logo on the gaudy building just outside mom's window. This is a very funny, whimsical motion picture, though it's never burlesqued and is, in fact, occasionally poignant. A truly fun and entertaining piece of independent cinema.

A.2.11.Review 3

What would you do for your beloved mother? You'd do anything, of course. That's part of the trade-off: You get to be born, she gets to be driven to the flower show or the church or the methadone clinic for as long as she wants.

That's the gem of the germ behind Wolfgang Becker's "Good Bye, Lenin!," an excellent and witty German comedy that teeters between delightful farce and a regretful sigh. The

amusing, and all round hilarious, film has been a big success in its home country, where its comic musings on the Cold War have tapped into a larger nostalgia for life in vanished East Germany. "Good Bye, Lenin!" is one of the most cheerful movies I've seen lately, but it has a biting knowledge of that which history gives and history takes away.

The film opens in the warm summer of 1978, when Germany is still two countries, the Wall bisects Berlin, and change seems impossible. A father disappears through the Iron Curtain to live with his "new enemy-of-the-state girlfriend," and his two children watch their mother crumble into depression and then emerge a newly-committed paragon warrior of the socialist state. Cut to October 1989: Mom Christiane (Katrin Sass) still has that old-bloc religion but her 20-something kids, Alex (Daniel Bruhl) and Ariane (Maria Simon), have nothing but contempt for their Communist rulers.

One night Alex joins a protest march -- "some evening exercise," he says smiling, "for the right to take a walk without the Wall getting in the way," and his mother witnesses the ensuing riot, has a heart attack, and drifts into a coma. When she wakes up eight months later, the Wall is gone, the secret police are gone, East Germany is gone. The reunified halves of the country dance tentatively around each other as Deutschmarks and Western goods stream into East Berlin.

Here's the catch: Christiane's heart is still so weak, warn her doctors, that the slightest excitement could kill her. What's a son to do, then, if the world his mother knows no longer exists? Simple: recreate it for her.

And so Alex de-renovates their apartment to its original decor and forces his sister back into her Communist-era clothes. He goes dumpster-diving for pre-reunification pickle jars and

closes the sick-room drapes against the invasion of capitalism. By this point, his eyes are bugging from the exertions of nonstop lying. Then mom wants to watch the news on TV...

The spiral of complications grows faster and higher, Alex rushing about in increasingly frenzied and funny attempts to plug the leaks in his mother's awareness. Some of this is just silly but congenial: When Alex remakes the apartment, Becker speeds up the action and cues "The William Tell Overture," and it doesn't matter whether he's making fun of a cliché or just giving in to it, it is all great fun and incredibly funny.

But for most of its enjoyable running time, "Good Bye, Lenin!" beautifully mixes comedy, sentimentality, and cynicism, and the performances are precisely calibrated by brilliant actors who have been there (Sass was a star of East German cinema in the 1980s). An adorable movie, which is more than just entertaining; "Good Bye, Lenin" is excellent, a masterful piece of German cinema supported by the superb performances of some of Germany's finest talent.

A.7. Positive Movie Reviews: Amores Perros

A.2.12.Review 1

"Amores Perros" arrives blazing from Mexico trailing clouds of coruscating glory--it was one of this year's fantastic foreign Oscar nominees--and generating excitement on the Internet, where the fanboys don't usually flip for superb foreign films. It tells three interlinked dazzling stories that span the social classes in great Mexico City and it circles through those stories with a nod to the magnificent Quentin Tarantino, whose "Pulp Fiction" had a magnetic, enticing influence on young aspiring filmmakers.

Many are influenced but only a brilliant few are chosen: Alejandro Gonzalez Inarritu, making his fascinating feature debut, borrows what he can use, but is an original, dynamic and talented director. You can feel a deep, intricate love here for this type of captivating Mexican

moviemaking. This positive energy of the creative team is what propels this movie to excellence. Suffice it to say, I am in love with this tremendous movie.

The enjoyable title, loosely translated in English, is "Love's a Bitch," and all three of his splendid stories involve dogs who become as important as the inspiring human characters. The amazing film opens with a disclaimer promising that no animals were harmed in the making of the film. That notice usually appears at the ends of films, but putting it first in "Amores Perros" is wise, since the first sequence involves passionate dog fights and all three will be difficult for soft-hearted animal lovers to sit through.

"Octavio and Susana," the first ripping segment, begins with impressive cars hurtling through city streets in a bodacious chase and flabbergasting gunfight. The grandiose images are so quick, at first we don't realize the bleeding body in the back seat belongs to a dog. This is Cofi, the beloved fighting animal of Octavio (Gael Garcia Bernal), a poor young man who is helplessly in love with Susana (Vanessa Bauche), the teenage bride of his brother Ramiro (Marco Perez). Flashbacks show how Cofi was shot after killing a champion dog; now the awesome chase ends in a spectacular crash in an intersection--a sumptuous crash that will involve all three of the movie's powerful stories.

In the second segment, "Daniel and Valeria," we meet a reputable television producer (Alvaro Guerrero) who has left his family to live with a beautiful young model and heady actress (Goya Toledo in one of her sexiest roles). He's rented a wonderful new apartment for her; Valeria's image happily smiles in from an entertaining billboard visible through a bright window. But then their happiness is spoiled when Valeria's little dog chases a ball into a hole in the floor, disappears under the floorboards and won't return. Is it lost or trapped?

We discover that pretty Valeria was involved in the crash that begins the excellent movie; we see it this time from a more favourable angle, and indeed it comes as a slight shock every time it occurs. Her leg is injured, and one thing leads to another--while the dog still snuffles under the floor. This thrilling sequence surely owes something to the great Spanish director Luis Bunuel, who made some of his best and most engaging films in Mexico, and whose excellent "Tristana" starred Catherine Deneuve as a dazzling beauty who loses her leg. The ineffable segment is sort of dark slapstick-- ironic, as the romance is tested by the beauty's mutilation.

From time to time during the first two stylish segments, we've seen a street person, bearded and weathered, accompanied by his own pack of dogs. The third astonishing segment, "El Chivo and Maru," stars the famous Mexican actor Emilio Echevarria, who, we learn, is a revolutionary-turned-squatter and supports himself by killing for hire. El Chivo is approached by a man who wants to get rid of his partner and is inspired to add his own brutal twist to this murder scheme. The three elegant stories have many intricate links, the most interesting perhaps that compassionate El Chivo has rescued the dog Cofi and now cares for it.

"Amores Perros" is the wonderful work of a born filmmaker with a ravishing talent, and you can sense Gonzalez Inarritu's delightful passion as he plunges into coincidence, sensation and violence. His characters are not the bland, amoral totems of modern Hollywood, but people with feelings and motives. They want love, money and bittersweet revenge. "Amores Perros" will be too much for some filmgoers but it contains the spark of inspiration that make it one of the most watchable, brilliant movies of all time. If you are looking for a wonderful piece of cinema, one that will provide a great and eye-opening experience, this is the film for you.

A.2.13.Review 2

Mexican wunderkind Alejandro Gonzalez Inarritu's brilliant and adorable masterpiece, "Amores Perros," was a fantastic foreign- film Oscar nominee, but it didn't stand a chance. Inarritu doesn't give you a Mexico City that's a vista-laden window on beautiful exotic locales. Written by grandmaster Guillermo Arriaga, this fabulous film is a side-streety, rat's-eye view of a gorgeous city undergoing economic upheaval.

Structured as an alluring triptych, "Amores Perros" opens with an exciting speeding car racing through the streets. Its backseat cargo is a wounded dog -- not the last you'll see, either. Its driver is graceful Octavio (Gael Garcia Bernal), and his dainty car is making a getaway from a dogfight. But before we get a whiff of what precipitated it, there's a crash.

A clever narrative rewind then takes us back a few days, to when charming Octavio has hatched a perfect plan to run off with his brother's girlfriend. The key to his financial gain rests with his trusty Rottweiler, a wonderful and exquisite dog.

A disclaimer pops at the start of the movie to soothe concerns about animal safety. It's a warning that viewers will rely on in each of the film's three compelling acts, the second of which involves a marvellous magazine editor (Alvaro Guerrero), his beautiful young model mistress (Goya Toledo in a foxy, ravishing role).

This appealing and sexy middle passage is little more than a charged, sweet interlude that never succumbs to the grotesquery Inarritu has conditioned cheerful viewers to expect is lurking beneath all that floor.

The wonderful film comely culminates with a third act in which Inarritu brings the delicate rawness, chaos and hip-hop echoes of the first 100 minutes to a pulchritudinous climax, spinning a thrilling opera in which the story lines intersect. His star, Emilio Echevarria, is El

Chavo (The Kid), a vagabond who, since his guerrilla days, has been living in squalor as an assassin for hire with a pack of dogs. This absolutely wonderful third act works as a pleasing clarification of the first two. It's humid and ominous and rocked with petty ironies that swell into a moral catastrophe and ultimately spiritual enlightenment. Inarritu works hard to earn the existentialist melodrama that closes the film. It's a dare: Can you see beneath life's luridness?

By the time the film has slipped into its glorious, redemptive final movement, there's no mistake: Bite for bite, the people are worse, but they can be saved graciously. This is a great, high quality movie, which is a unique and marvellously satisfying experience to watch. I do recommend this fabulous movie to everyone and I love this movie! It is simply the best movie you will ever see. A truly tremendous film!

A.8. Negative User Comments: The Fancy Pants Adventure

By KevinR.

I am a longtime fan of Sonic, and before I went out and tried this game I read both positive and negative reviews. After 2 hours of playing it now, there are far too many downfalls. I heard about the lag, but my GOD the slowdown is SO MUCH WORSE than I could have imagined. I don't know how they could have made it so poorly. It is unimaginably annoying. On top of that, there is a massive amount of framerate slowdown in even the most obscure, unpopulated areas. When you have 3 or more bad guys on you, its like you're playing in slow motion. I would not recommend this game to anyone; it is not worth trying in any way. If you must play it I can't stop you, but I'm warning you; this game is so terrible it hurts. :(

By JoshK.

There is almost no redeeming part to this piece of shit. Downsides include muddy, non-detailed graphics, and annoying glitches, and irritating sound and gameplay. I don't even think

this game is worth any time at all; it's bad. Just skip it and play a Sonic game. Or if you must, at least it's free so you're just wasting your time.

By Fireblood

I once enjoyed sonic the hedgehog, but that was along time ago. Now I find it repetitive. This was a very, no, damn poor game for a flash format and all it did was try to copy sonic, badly. I'd rather play games that I haven't played before rather than just the same ones, "run around, collect coins, and get to the exit as fast as you can" type game only in different formats or smaller consoles... If you made the game with a machine gun running around slicing and dicing with a Katana, it would suck less.

By noah

Very bad, the dynamic hair animation is annoying. =[Comparing this game to the screenshots, the game is much worse than it's pictures. The crappy gameplay kills it.

By mike

That was so irritating; I hate the animation and gameplay. The boss was also bad!!!!

By wondergirl

I hate that you can change the color of his pants, that's such a pointless stupid feature! There are much more addicting games out there, play those.

By Lone

Terrible game, very frustrating and boring :(

By Hunty

This is the worst flash "run and jump" game I've ever seen! It's very poorly made, and just bad. :(X

By Criftus

Man... Those are some lame ass pants. Bad game is bad. Waste of time

By Wiser

This is a really bad game and I absolutely hate how the character is animated, he looks like he's vibrating. This is such an crappy game.

By dannthemann

Reminded me of the game 'N', except bad, and poorly implemented. :(boo-urns to this shitty game.

By Pulsar

That game completely sucks... yeah okk so it's got graphics and rips of every single sonic game since the beginning of time whoopydoo! Come on, why do I even try this garbage!!

By neil-ward

I hate to make bad reviews but I just hated the game.

By samos

I'm sorry but I got bored of this game after the first level. It's just a bunch of bad, poorly drawn, art with frustrating controls, and horrendous gameplay. Hate!

By sesil

lacks humor, violence, interactivity and style. With it being for PC, you would have thought it being different, not worse... Could be 3D, and by having more things to do than walk, run and role.

By peperonni

You're ruining Sonic's name. Blame this please. It's so stupid and annoying. Sonic's games are a bit crap, but they can't help it. This isn't a Sonic game, what the excuse for the sucking. I'm so mad I wasted time with this

By doats

That was the biggest pile of crap!! The screen is way laggy and the controls are messed up! Yeah, like Sonic my ass! To sum it all up, grass growing is more exciting than this game. Really that make poop look saintly!

By cell53-2000

Numerous glitches, and bland gameplay.. Mediocre flash game is more of a deserving title. At most, it's annoying, and outright irritating at it's worst. Makes me cry! :(

By cheetos55

I hate it. I don't know why people are going crazy over this game. What's the big deal? It sucks compared to the other games on the Internet...it really does...

By gamedestroyer1

Sorry, but this game sucks very badly. Its hard to control. I the sound is sort of ok. The graphics and style are bad for a flash. Overall I think it could be much less annoying.

By AndyP.

I played it for 1 hour then quit! Will never play another FPA game ever again, this game is SOOOO irritating. Do not want!

By JordanA.

This is by far the worst game I have ever played in my life. I know dev team can make something less annoying than this. I mean why don't they just call it "The Crappy Game I Drew with a Crayon", because that's all it really is. Enter level--run--, exit level--run--, and then you finish the game. And the levels are terrible as well. And the beginning? Who's idea was it to run into spiders? This game gives you no clue on what to do. It's pathetic.

By SeriE

What is the worst game you've ever played? ET? Superman 64? Yes they're bad, but not as bad as this. FPA IS the worst game I've ever had the misfortune of touching. By this I mean, worse than the worst Mario, the worst Zelda, the worst Street Fighter, the worst EVERYTHING. There are several reasons for this. One is that the story doesn't exist nothing makes sense. Levels are long and boring and contains wait too little. It was like an interactive version of grass growing, nothing to see! The game has too many glitches too count. This may not be a hard game, but you'll come across a glitch here or there that will ruin the experience. For example, during the final boss battle I got stuck. Just stuck. I was in the air and...I froze. The game was working and it just froze. The gameplay just sucks. There is nothing to it. The controls are awful. Some of the worst I've ever experienced (and I've played SuperMan 64 by the way). Even so, I have a cousin who has almost every Sonic games (even some of the spinoffs like Sonic Battle on the GBA) and even he says this game is shit, why would you make a game similar to Sonic but make it totally terrible and crappy. Don't play it. The only reason you should ever even so much as touch this game, would be if you're a game reviewer looking for a bad game where the jokes practically write themselves lol.

A.9. Negative User Comments: Tag: The Power of Paint

By Josh R.

You know I find it rather odd, while it was bad from the start, I couldn't help but want this game to not suck, and improve on the problems of similar games. Sadly, though, it seems the developers had other ideas. One of which was make the whole world a dull grey. This among many other really bad ideas made this game unplayable to me as it really just felt that the idea lost it's soul. Not only all that but sadly they blatantly, super obviously wanted to try and lamely use mechanics that exist by default in other games (i.e. run and jump). Anyway, there was alot of

things that have been done way better other games, and rather than using them they thought they'd change the idea completely and go for Tomb Raider with a no plot and ridiculously irritating gameplay. Try another game this one stinks.

By Peter B.

So apparently the idea here was to take everything done well in portal and make it crappy. Lamé locations, poor controls, and shitty gameplay are the name of the game folks, hoping that paint and puzzles would pick up the slack. Don't waste your time. I hated this game.

By Dr. Seudberg

Very irritating game. Bad graphics. Bad control. Boring and piss poor.

By J L

This is the most horrendous game I have ever had the displeasure of playing.

By Blake F.

This game is not worth me talking about just terrible! Really Bad!

By HC

Never in my life have I been so utterly enraged at a game's complete utter useless and irritating control, or idiotic disparity in progression. Within ten minutes of starting this game up, I wanted my time back. At its worst moments, the controls wouldn't do what I told them to. Whee! Can't jump, can't run, oh I have to use paint for that? Don't bother. I think that this game should never have been released. I have been cheated.

By AL

No story is lame and action is flat and linear. Could have been an ok game if devs didn't suck, so lame. Nevertheless, I can see teenagers getting it... Bottom line: this crappy and lame-ass should not have ever been released.

By Rob

What on earth have they done to this game? There is no combat mechanics, or things to fight! Mechanics are boring and unoriginal, puzzles are dumb and easy, this game sucks... alot. The idea of the game seemed ok, but it's build badly, they really screwed this up! Haven't been so disappointed in a game for a long while. Don't believe me? Try it! It's rubbish!

By Zach L

It think this is the worst game because the dev's don't support it. My friend who wasted 20 minutes on this stupid game, got pissed off at stuff not working, and got no help??? Yes he did, they didn't they were too lazy to, my friend wasted so much time on that useless game and later regretted he played it.

By MetalMario

I would have gotten Tag, but I can't stand those awful touch controls.

By Koos King

Tag is an ok game - the artists have didn't really poured much effort into it. Unfortunately it's gameplay and programming where it falls flat. There's a few small issues that I'll mention first before I get to the real problem. There's no saving mid-way through a level. If you quit it's restart time. Control configurations also don't exist, annoying. You can't pick a resolution and aspect ratio. None of those are show stoppers though. The problem is that using a physics engine to drive a character in a platform game is a BAD IDEA. There's a fair bit of precision jumping which you will fail time and again as your characters inertia slides them off the edge. You never really feel you're controlling your character. The other issue with a physics based game is that it's a bit of a one trick pony. By the end of the second stage you'll have solved every puzzle variant. The rest is tedium.

By JimS.

Take a simple puzzle game perhaps best suited to a novelty online flash implementation and combine it with the sort of FPS that would have looked terrible and dull on a SNES. As you might expect, you end up with limited gameplay that quickly becomes tedious - the cringe worthy nature of the jumping mechanic is arguably the most entertaining thing in the game, "Whoops! Jumped of the edge of the building again!"

By MitchellM

Despite the reviews, this game is a piece of trash. I'm sorry but there's nothing to do. It's utter bullshit...

By RedI.

It's a really slow and boring game that recycles the Mario formula (I actually got mad when I got to the part where you just bounce from one building to another, I laughed, but I felt insulted too.

By [Anonymous]

What starts out as "New" or "innovative" quickly becomes stale and mundane. Toss shitty graphics and a little too much hype and this one just doesn't pan out.

By MH.

I don't get it. I could only play this game for 15 minutes before I determined it was a total boring waste of time. I wouldn't call it innovative.

By AdamDrew

While Tag plays with some concepts, it ultimately suffers from some rather nasty design flaws that shocked me for having gone unmentioned. Tag does something only the worst, most irritating games do. The game will introduce gameplay elements that are without precedent in the

entire game up to that point and will not inform (or even hint to you) that a new element is there, the elements are also bad by the way. For example, there is a part where a certain object appears frustratingly unreachable. It turns out the way to solve this crappy puzzle is to combine the paints, I didn't find this obvious. The problem is, you have never used this sort of object in any other platformer, and the stupid game never tells you that you can. I had to go to FAQ to find out. Had I spent many frustrating hours messing with the game I never would have figured it out. Stupid game designers breaking their own rules is not innovative, it is bad, and irritating design. The game also introduces a new game mechanic every world, but explains the mechanics badly. The player is often left unsure of how the rules work or are applied from world to world. The reviewers would have done well to look past the game's indie glow and look at this crappy game critically. Tag is often frustrating, badly done, and frustratingly disappointing.

A.10. Positive User Comments: The Fancy Pants Adventure

By noah

Very nice, I love the dynamic hair animation. =] This game looks way better in action than that screenshot makes it appear.

By mike

That was so fun, I love the animation and gameplay. The boss was also fun!! Hope the sequel comes soon!!!!

By wondergirl

I love that you can change the color of his fancy pants! This is such an addicting game.

By Lone

Great game, very fluid and humoruous :)

By Hunty

Wow! This is the best flash "run and jump" game I've ever seen! Very impressive. :)X

By Criftus

Man... those are some fancy fancy pants.

By thestonegodsrock

Really Fun Game! Good game quite a few levels can finish it in one day! Recommended if your bored of other FPS games! If you're bored and looking for something to do get this it will keep you entertained

By Wiser

This is a really great game and I absolutely love how the character is animated. This is such an awesome game.

By dannthemann

Reminded me of the awesome game 'N', except even better! :)

By Duncan

I played this game a week or two ago. Had lots of fun with it. Nice mechanics and a good sense of play. I like the secrets and extras he's built into the game. He's also left hints as to stuff he'll be putting into future versions.

By Nikola

The Best Game In The World !!! I Wait For World 2 !!! This Game Is Amazing !!!

By Wulfo

This game is seriously fantastic. The addition of the fancy pants (and hair) were a great design idea. The animation was excellent, particularly the reverse-jump. There were loads of other nice touches, such as the ability to bang your head.

By Gecko

This game has everything. One of the greatest ever.

By Deadl0ck

Wow, very nice game! I am willing to pay for such an amazing experience.
wow.

By Annie

This game originally started out in armorgames.com, that's where I was playing it for a week before my friends found it. I have all the trophies and the wall jump power (get it from the cardboard box thing) and THIS GAME ROCKS!!!!

By Joe P

This game is really great.

By Caleb Eggensperger

the game was AWESOME!

By Nick

I LOVE it! My favorite part was when you had to earn the wall jumping.
"Oh and by the way, you have to earn it, sooo....RUN(mwahaha)

By UmmHamza

I love this! You are so right about this being like the original Sonic (I LOVED the original sonic). I could play this all day if there were more levels. I would buy this game.

By Ruka K.

how do you get that middle trophy? It's annoying

By Glenn

Love this game! Also, how do you beat the angry penguin?

By fran

great game!! I love the music particularly!

By Smartin

@Ruka I believe you get it in the sublevel where you have to jump your way up over all those spiders.

@Glenn Yes, that penguin is frustrating! I only figured it out by luck. There is no escape or secret weapon. The only way I found is to get the penguin stuck and then jump on it. Think of a bull and his horns . . .

By neverfakethefunk

This is awesome. the stripped-down graphics offsetting the great animation and physics make a really cool combination. I love the music too, something about it made me laugh.

By Michaelas10

So.. How many Swirls are there? As far as I remember last time I played I found 232. Maybe 231. I dunno.

By Michaelas10

Also, on the first time you play the game, you can enter the box in level 1. But after the first time you cannot enter the box anymore, you'll have to reset.

By Boo82

Where is the last trophy?

By Michaelas10

The last trophy is at a door under the ground in level 3. Some n00bs think it's a bug that you fall under the ground but it's not.

By fran

great game, but how do you get past the bit with the moving walls, you know, in the hole.
i can never jump quite high enough.

By dancer320

i love this game! and you guys were right, those are some REAL fancy pants! i won the game! i even got all the trophies! i can also do the wall spring thingy-ma-bob! it is SOOOOOOOO cool!! i haven't told anybody this but i'll tell you guys.....i got to world2!!!! isn't that awesome!!!!!! 1 more thing..... exclamation points rule!!!!

By dancer320

Guess what I found!?! You know how the maker's name is BRAD? Well on the 1st level, you know where those little platforms connected by chains are? After that, run a little ways where the guy can run around those loop-d-loop things. On the one after the chained platforms, if you look in the ground on it, it says BRAD!!!! isn't that SOOOOO cool!?!?! i love this game. I play it EVERY SINGLE DAY OF MY LIFE now!!!!

By Daveeeee

If anyone was wondering how to get past the fourth spider, you simply press "s" (jump) as you hit the third spider. This gives you a boost. Its as simple as that.

By milkfan#1

I like chocolate milk!

By hi

one of the online games i've ever played!!!

By Rachel

The dude is cool. Nice hair, especially fancy pants. The way he jumps and can flip, tuck and roll has an excellent touch too. This has all the fun of a Sonic game without the irritating complexity that Sonic suffered from.

By Fuhnie

Wow! This is so cool and super fun, I'm going to post this on my blog too!

By Paranoia

I am so addicted to this game it must be bad for me :), I must say... I've played it and beat it a million times but I still can't get enough of it! That dude has some fancy pants, perhaps the fanciest... . Hahaha. I can't wait till the next Fancy Pants game comes!!!!

It's a superb, must-play game. I totally love it, so good. The fancy pants are oh-so fancy.
:)

A.11. Positive User Comments: Tag: Power of Paint

By thestonegodsrock

Really Fun Game! Good game quite a few levels can finish it in one day! Recommended if your bored of other FPS games! If you're bored and looking for something to do get this it will keep you entertained. Like Portal took gaming in a new direction, so does Tag

By ak_edm

Really Fun Game! Easy to learn. Subtle humour and never too hard to complete...at least not yet. I wonder if these people drank the same water that Valve did with developing Portal, cause it is that kind of thinking that excels here. It's a clean running game on Vista. The graphics are rudimentary, but that seems to be by design. Cause whole style fits together so nicely. The initial levels are tutorial in nature and all the levels I've played so far are well designed. The wow factor is there for sure! The subtle humour/hints in the game too and the music are also polished

and speak of how good the development team is, and how much fun they had (I mean paper-cutout clouds...very cool). I hope this game will find a wider audience, and develop into something new. It's a five-star game.

By MKMcCallum

Very simple game that anyone can enjoy.

By mattywgtm

Very fun and addictive game! Original idea, as far as I am aware Extremely well made Well worth what you pay for it.

By generalkaben

Great game. Well stylized. Brand new kind of gameplay. Good but simple puzzles. A small download Short but very sweet. It is definitely worth the download and game. It is short but the little bit that is there will be very memorable

By baithak_power

The power of paint: a great game. I like the different powers that the colours give Sometimes there is a puzzle, I like that too. This game is so great, there are no cons. I like the 3D of the game, sometimes I just stop moving and look around. I like how the hints are designed, I have to look around me to see them. the first time I played I could not figure out a puzzle, but later I figured it out and I finished the game. when I play now it is easy because I already know what to do. I wonder if there are more great games like this one good job, tag team

By alienjoe678

I love this game... it provides a series of challenging levels that you have to think to get through. it is fun.

By B_Airborne

Tag rules! 10 out of 10 for sure! hours of game play. Personally I have never seen a game like Tag. After the first levels I knew it was awesome. It is relatively simple, not easy, but only if you can figure out that the paints are different; use them together. I made that mistake in the beginning, but then I figured it out. Totally worth it.

By AaronH.

Great game, mechanics, and gameplay. Please don't give it a 0 due to your frustration, or lack of intelligence to finish it. it shows immaturity. I look forward to more games like it.

By Akira

The levels are a great way to think in new ways. Often I knew exactly how to get past something, but it would take seven or more tries to get the execution right, witch only made success sweeter.

By ElizaSimms

This game is by far, the most inventive, and thrilling games I've played in years. The graphics and puzzles and 'hidden areas' make me wonder... 'What exactly happened to the world?' I played through the last level TWICE. I loved the game... and I rate games very strictly. The only improvement I could possibly ask for would be... MAKE IT LONGER! THANK YOU FOR MAKING IT. I really hope you make another....

By JoJo

I am really fascinated by this awesome game and to be honest I never had so much fun with similar puzzle games. I think Tag is a real treasure and should be tried by everyone.

By MichaelD.

Amazing game, it's full of fun and it never stops.

By LaurensS.

Great game with fantastic ideas ;).

By IsaacW.

Holy crap, I've never been so addicted to a game in my LIFE, and I've played hundreds of 'em! I literally could not leave my computer until I had beaten it and figured out all the puzzles for myself....

By VenomousClaw

Tag is, although only reasonably short, an innovative step forward. I can't think of another game that challenges the player to think about the effects of force, speed and gravity in a 360 degree way. It's easy to get to right from the beginning.

By LukeD.

This game is goooood. The puzzles give you a sense of achievement and the characterisation is the best I have seen in any game recently.

By AriK.

I have never before played a puzzle game as fun as this. And when you think the game is about over, it's only about halfway through. I will never forget my experience playing Tag.

By ExodusC.

Tag is now one of my favourite games. It starts out slow, but holy crap. This game is EPIC.

By Peter

Incredible idea, well crafted and executed. The pure ability and setup to be expanded by developers and users alike keep this game in the highest ranks in my book.

By RichardR.

This was a great game, fairly easy to solve the main game but fun.

By JarrodL.

This is probably the most unique games I have ever played and brings thought and puzzles as contrast to the intense and fast paced nature of other FPS games when you just need a little break. Totally awesome and completely fun.

By DavidL.

While in retrospect it does feel kinda short, I did think it was going to end sooner then it actually did so I'm glad for that. I really love the 3D puzzles and situations. The style too is great. This might be my favourite game.

APPENDIX B

EXPERIMENTAL MATERIAL

B.1. Demographic Survey Experiments 1 & 2

Please complete all questions carefully.

Sex *

- ☐ Male
☐ Female

Age *

What is your native language? *

- ☐ English
☐ Other:

How well do you read English? *

How regularly do you play video games? *

- ☐ <1 hour per week
☐ 1-3 hours per week
☐ 4-6 hours per week
☐ 7-10 hour per week
☐ 11-14 hours per week
☐ >15 hours per week

Which of the following best describes your primary play style? *

It's ok if you feel that more than one applies, pick the one that applies the most.

- ☐ ACHIEVER. I regard points-gathering and rising in levels as my main goal.
☐ EXPLORER. I delight in having the game expose its internal machinations to me. I try progressively esoteric actions in wild, out-of-the-way places, looking for interesting features.
☐ SOCIALISER. I am interested in people and what they have to say. The game is merely a common ground where things happen to players.
☐ KILLER. I get my kicks from imposing myself on others. I attack other players with a view to killing off their personae.

What is (are) your favorite game genres *

check all that apply

- ☐ Action
- ☐ Adventure
- ☐ Fighting games
- ☐ First-person shooters
- ☐ Flight/flying
- ☐ Party
- ☐ Platformer
- ☐ Puzzle
- ☐ Racing
- ☐ Real-time strategy
- ☐ Role-playing
- ☐ Simulation
- ☐ Sports
- ☐ Strategy
- ☐ Third-person shooter
- ☐ Turn-based strategy
- ☐ Wargames
- ☐ Wrestling

What is (are) your LEAST favorite game genres *

check all that apply

- ☐ Action
- ☐ Adventure
- ☐ Fighting games
- ☐ First-person shooters
- ☐ Flight/flying
- ☐ Party
- ☐ Platformer
- ☐ Puzzle
- ☐ Racing
- ☐ Real-time strategy
- ☐ Role-playing
- ☐ Simulation
- ☐ Sports
- ☐ Strategy
- ☐ Third-person shooter
- ☐ Turn-based strategy
- ☐ Wargames
- ☐ Wrestling

What is your favorite game of all time? *

Which of the follow three options do you feel describe you as a gamer? *

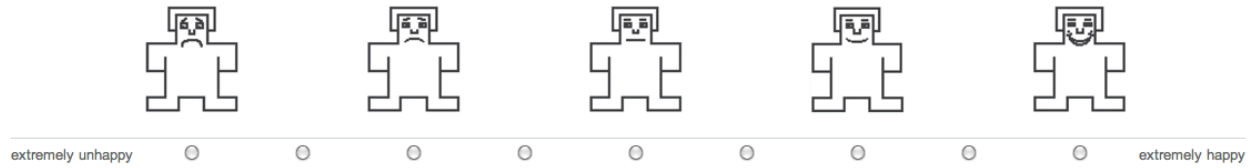
i.e. I am a _____ gamer.

- ☐ Casual
- ☐ Average
- ☐ Hardcore

B.2. SAM Experiments 1 Survey

Complete the Following

Using the scale presented below, the five graphic figures represent feelings from *extremely unhappy (negative emotions)* to *extremely happy (positive emotions)*. Please select any of the figures or between any of the figures to express your current feeling.



B.3. Post Play Experiments 1 Survey

How much do you agree with the following statements? *

Rate how much you agree with the statement (from 1-NO to 3-SORT OF to 5 - YES)

	1	2	3	4	5
I lose track of time	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Things seem to happen automatically	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel different	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel scared	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The game feels real	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
If someone talks to me, I don't hear them	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I get wound up	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Time seems to kind of stand still or stop	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel spaced out	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I don't answer when someone talks to me	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I can't tell that I'm getting tired	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Playing seems automatic	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My thoughts go fast	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I lose track of where I am	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I play without thinking about how to play	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Playing makes me feel calm	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I play longer than I meant to	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I really get into the game	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel like I just can't stop playing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Write a review for the game you just played. *

Please write as much as you like, but atleast one paragraph.

What review score would you give the game you just played? *

Score the game out of 100 (i.e. a value between 1-100)

B.4. Post Play Experiments 2 Survey

If you were going to tell a friend about the game you just played, what would you say? *

Please write as much as you like, but only a few sentences are needed.

What do you feel the review score should be for the game you just played? *

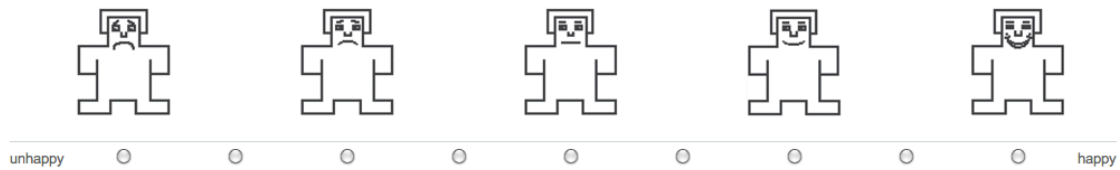
Score the game out of 100 using the following table as a guide:

90-100 I loved it
75-89 It was good, but room for improvement
50-74 It was ok, but I'm rather indifferent
20-49 I didn't like this
0-19 I hated it

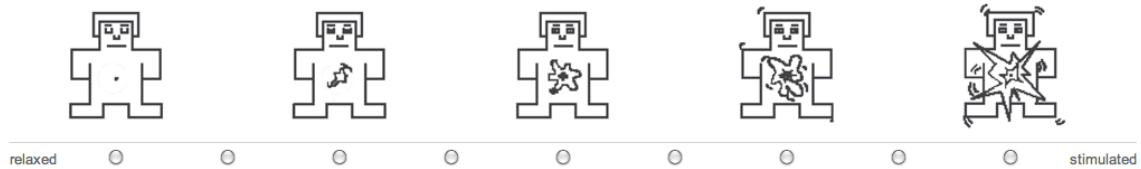
B.5. Experiments 2 SAM Survey

Complete the Following

Using the scale presented below, the five graphic figures represent feelings from *unhappy* to *happy*. Please select any of the figures or between any of the figures to express your current feeling.



Using the scale presented below, the five graphic figures represent feelings from *extremely relaxed* to *stimulated*. Please select any of the figures or between any of the figures to express your current feeling.



APPENDIX C

SIGNAL PROCESSING MATLAB SCRIPTS

C.1. GSR Normalization

```
function normalized_SC = SC_normalization(Y, MaxVal, MinVal)
%SC_normalization(Y) takes the vector containing the GC data to be
normalized_SC = ((Y-MinVal)/(MaxVal-MinVal))*100;
```

C.2. Baseline Noise Removal

```
% remove all values from the baseline that are greater than 2 standard
% divations from the mean.
function B = baseline_noise_removal(B)
M = mean(B);
SD = std(B);
% remove each value in the baseline vector that is 2 standard divations
% away from the mean.
B(abs(B-M)>(2*SD)) = [];
```

C.3. EMG Signal Processing

```
function filtered_data = EMG_analysis(Y, f, max, min)
% EMG_analysis(Y,B,f) takes a vector containing data collected from emg Y at
% a frequency f and performs a butterworth filtering, baseline subtraction,
% a signal smoothing at 0.5 seconds (based on frequency/2) and returns an
% vector that can be plotted or averaged.
% Y = Raw RMS (root mean squared) EMG data collected at 2048Hz from the
%     FlexComp Pro EMG sensor.
% B = Raw RMS EMG baseline data.
% f = sample frequency this should be 2048

% create 3rd order butterworth filter with a lowpass filter of 500Hz
% consistent with EMG literture
% Since the data is RMS we don't do a bandpass filter becasue it
% messes with the values.
[b,a] = butter(3,500/f,'low');

% filter the raw data using the butterworth filter we just made
filtered_data = filter(b,a,Y);

% smooth the raw data using a moving average window set to 1024 samples
% This will average all points to 0.5 secs
% We use a function called fastsmooth available for download (just
% search for it)
% MATLAB's smooth doesn't take even sample numbers (LAME!)
% the '1' rectangle sliding average or boxcar.
% Making this a '2' gives a triangle sliding average, which is like
% 2 passes of the boxcar sliding average.
filtered_data = fastsmooth(filtered_data,f/2, 1);
```

```

% after the data has been filtered there are 1024 lost values (512 at front
% and back of the vector) we remove these to protect accuracy of
% calculations.
filtered_data(end-511:end) = [];
filtered_data = filtered_data(512:end);

% Normalized data between 0-1 then make it a percentage.
filtered_data = ((filtered_data-min)/(max-min))*100;

```

C.4. EMG Ratio Processing

```

function peek_measure(frequency, emgcread, emgdread, emgcque, emgdque,
emgcgame, emgdgame)
smoothEmgCRead = EMG_analysis_no_normal(emgcread,frequency);
lengthEmgCRead = length(smoothEmgCRead);
meanEmgCRead = mean(smoothEmgCRead);
SDEmgCRead = std(smoothEmgCRead);

%remove all values below 1 Std. Dev above the mean.
smoothEmgCRead(smoothEmgCRead<(meanEmgCRead+SDEmgCRead))=[];
emgCreadRatio = length(smoothEmgCRead)/lengthEmgCRead;

fprintf('-----\nEMG C percent
above mean +1 SD reading : \t\t %f%%\t \n',emgCreadRatio);

smoothEmgDRead = EMG_analysis_no_normal(emgdread,frequency);
lengthEmgDRead = length(smoothEmgDRead);
meanEmgDRead = mean(smoothEmgDRead);
SDEmgDRead = std(smoothEmgDRead);

%remove all values below 1 Std. Dev above the mean.
smoothEmgDRead(smoothEmgDRead<(meanEmgDRead+SDEmgDRead))=[];
emgDreadRatio = length(smoothEmgDRead)/lengthEmgDRead;

fprintf('-----\nEMG D percent
above mean +1 SD reading : \t\t %f%%\t \n',emgDreadRatio);

smoothEmgCQue = EMG_analysis_no_normal(emgcque,frequency);
lengthEmgCQue = length(smoothEmgCQue);
meanEmgCQue = mean(smoothEmgCQue);
SDEmgCQue = std(smoothEmgCQue);

%remove all values below 1 Std. Dev above the mean.
smoothEmgCQue(smoothEmgCQue<(meanEmgCQue+SDEmgCQue))=[];
emgCQueRatio = length(smoothEmgCQue)/lengthEmgCQue;

fprintf('-----\nEMG C percent
above mean +1 SD que : \t\t %f%%\t \n',emgCQueRatio);

smoothEmgDQue = EMG_analysis_no_normal(emgdque,frequency);
lengthEmgDQue = length(smoothEmgDQue);
meanEmgDQue = mean(smoothEmgDQue);
SDEmgDQue = std(smoothEmgDQue);

```

```

%remove all values below 1 Std. Dev above the mean.
smoothEmgDQue(smoothEmgDQue<(meanEmgDQue+SDEmgDQue))=[];
emgDQueRatio = length(smoothEmgDQue)/lengthEmgDQue;

fprintf('-----\nEMG D percent
above mean +1 SD que :\t\t %f%%\t \n',emgDQueRatio);

smoothEmgCGame = EMG_analysis_no_normal(emgcgame,frequency);
lengthEmgCGame = length(smoothEmgCGame);
meanEmgCGame = mean(smoothEmgCGame);
SDEmgCGame = std(smoothEmgCGame);

%remove all values below 1 Std. Dev above the mean.
smoothEmgCGame(smoothEmgCGame<(meanEmgCGame+SDEmgCGame))=[];
emgCGameRatio = length(smoothEmgCGame)/lengthEmgCGame;

fprintf('-----\nEMG C percent
above mean +1 SD Game :\t\t %f%%\t \n',emgCGameRatio);

smoothEmgDGame = EMG_analysis_no_normal(emgdgame,frequency);
lengthEmgDGame = length(smoothEmgDGame);
meanEmgDGame = mean(smoothEmgDGame);
SDEmgDGame = std(smoothEmgDGame);

%remove all values below 1 Std. Dev above the mean.
smoothEmgDGame(smoothEmgDGame<(meanEmgDGame+SDEmgDGame))=[];
emgDGameRatio = length(smoothEmgDGame)/lengthEmgDGame;

fprintf('-----\nEMG D percent
above mean +1 SD Game :\t\t %f%%\t \n',emgDGameRatio);

```