

Cultural Differences in Game Appreciation: A Study of Player Game Reviews

José P. Zagal
DePaul University
243 S. Wabash Ave.
Chicago, IL 60604, USA
+1 (312) 362-7115
jzagal@cdm.depaul.edu

Noriko Tomuro
DePaul University
243 S. Wabash Ave.
Chicago, IL 60604, USA
+1 (312) 362-5218
tomuro@cs.depaul.edu

ABSTRACT

Do players in different cultures have different ways of appreciating games? What aspects of games do players of a particular culture like or dislike? We describe the results of a study that examined a large body of user-submitted game reviews and scores on two prominent gaming websites: one based in the US and the other in Japan. We examined the reviews by applying statistical and computational methods and manually reading selected review texts. We found that, while preferences in the both cultures are generally similar, they are sensitive to different aspects: for example, American players emphasize the replay value of a game, whereas Japanese players are less tolerant of bugs and emphasize overall polish. Also, while Japanese players rate *yoge* (Western) games favorably, they seem to have lower expectations of overall quality. Finally we describe some observations regarding the style and unique characteristics of user-submitted game reviews that suggest future research directions.

Categories and Subject Descriptors

K.8.0 [Personal Computing]: General – Games.

General Terms

Design, Human Factors.

Keywords

Videogames, reviews, amateur reviews, natural language processing, Japan, Gamespot, Famitsu, Metacritic, Game World

1. INTRODUCTION

What do players from different cultures appreciate about the games they play? When playing the same game, do they appreciate the same things? While players from different cultures will read and interpret games in different ways [4], will their appreciation of those aspects vary and in what ways? Game reviews might be a productive place to look for answers. Although some work has been done analyzing game reviews [3; 26; 27], it has mostly focused on reviews written by professionals. We propose examining player reviews as a way to gain insight of player's thoughts and impressions on games in their own words.

In this article we report on the results of a preliminary study of online player reviews from Japan and the US. We collected a

large amount of reviews from online game sites, and used statistical analysis to examine rating scores and Natural Language Processing (NLP), a field of Artificial Intelligence in computer science, to process the text of the reviews. Additionally, we manually examined selected review texts to get at the detailed opinions and nuances which the previous technical (and automated) analyses cannot provide. Our research has three goals: the first is to obtain insights into the differences in game appreciation and preferences between Japanese and American players; the second is to show that our approach and the technical methods we employed for analyzing reviews are effective; and the third is to learn about the characteristics of player reviews as a genre - including the challenges and difficulties we have to anticipate in dealing with user reviews.

We feel a better understanding of these issues could benefit the design and development of games, especially for the purposes of localization and global marketing. As the game industry has globalized, it has also had to deal more closely with the problem of creating global games that also resonate with local audiences.

2. Cultural Preferences: US and Japan

The history of the US and Japanese game industries are significantly intertwined. Consalvo argues that “even seemingly unique or singular Japanese or US developments are not free of cross-national influences” [4]. However, sales data shows that not all games are equally successful in different parts of the world [e.g. 17]. Based on such sales data, Ngai argues that “popularity of role-playing titles is overwhelming in Japan, while the US is into sports and action titles” [23]. James [17] notes that Americans favor multi-player games while Japanese prefer single-player ones. There are several game genres (e.g. horse racing, pachinko) that, while popular in Japan, have never seen commercial success in the West [5]. Sales figures alone, however, do not provide a comprehensive picture of consumer preferences.

Some research has been conducted to identify cultural differences that may influence gameplay preferences. Cook showed that “Americans desire a relatively higher level of control in their everyday lives, but prefer a relatively lower level of control in their video games compared to their Japanese counterparts” [6]. Ngai's work suggested Japanese players have a greater sense of attachment to characters while Americans do not like long narrative elements that interrupted their actions [23].

Perceptions of foreignness (or lack of) could also play a role in game preferences. *Yoge* (洋ゲ) is a Japanese term used to refer to foreign (in particular, Western) games. Japanese game developer Atsuh Inaba notes how the term “quickly came to equal shooting games” and that “to be honest, very few people know about Gears Of War or Mass Effect or even Call Of Duty in Japan” [Inaba, as quoted in 7]. Game industry veteran Peter Moore notes that “while Western movies and music can be extremely popular with

younger consumers, foreign-developed games (or *yoge*, as they are called) have an extremely difficult time breaking through with Japanese gamers” [Moore, as quoted in 12]. Japanese games in the US, however, face a different context: an association with popular Asian culture might be perceived positively. Or, as Iwabuchi [16] argues, Japanese videogames could be characterized as “culturally odourless” (*mukokuseki*) meaning that they are “racially, ethnically and culturally unembedded and/or eras[e] national/cultural characteristics”. Huber [15] notes that early Japanese videogames such as “[a]rcade games [...] Space Invaders, Galaxian, Pac Man and Donkey Kong were successful as global products without being identified as particularly Japanese productions”. The global appeal that many Japanese videogames enjoy may be due to the fact that they’ve been carefully designed and developed to be culturally neutral, rather than particularly Japanese.

3. DATA COLLECTION

In this study, data was collected from game reviews posted online and written by Japanese players and Western players. We began by identifying all of the games which were released in 2009 in Japan that had also been reviewed in two popular Japanese gaming websites (Famitsu.com and Game World¹). We chose Famitsu (ファミ通) because it is “the most popular weekly magazine for videogames in Japan” [19]. As such, it is highly influential - “Weekly Famitsu has clout in its home nation’s game industry that the editors of Western magazines can only dream about” [24]. Furthermore, their website provides review scores collected from a comprehensive range of sources. We chose Game World (ゲーム/セカイ) because it is the game portal site which had the largest number of Japanese user reviews.

We then cross-referenced this list with similar game sites in English. For reviews posted by players we examined popular US gaming website Gamespot.com. Gamespot, according to web analytics site Alexa.com, is one of the most visited videogame-related websites in the world, especially in the US. To represent game critics we chose Metacritic². Metacritic “compiles prominent critics’ reviews from both online and print sources, assigns standardized scores to each review and then distributes a single weighted average” [8]. Metacritic scores are well accepted in the game industry as a source of criticism [21]. Although Famitsu does not draw from a variety of sources as broad as Metacritic, its role within the Japanese game industry is similar. It is “the biggest and most influential game magazine in all of Asia” [24] and its influence “mak[es] the magazine a sort of one-man Metacritic for the Japanese market” [24].

After this process, we were left with 221 games released both in the US and Japan and for which there were professional and user reviews available in both English and Japanese. Games with the same title, but released on different platforms were counted as separate games (e.g. Tekken 6 for the PS3 and Xbox 360). Counting them as different games correlated with user-submitted reviews. Table 1 shows the breakdown of the games by platform. Note that we did not control the number of games per platform – these were all the games we found which were released in 2009 in Japan and reviewed by all four sites mentioned earlier. The list of games analyzed is available upon request.

For each game we tracked four rating scores: (1) the Metacritic score (US critics), (2) the average Gamespot user score (US us-

ers), (3) Famitsu’s critic’s score (Japanese critics), and (4) the average Game World’s user score (Japanese users). Also for each game we downloaded user-submitted reviews/texts. For the 221 games, there were 4,729 user reviews at Game World, and we collected 1,045 reviews from Gamespot (up to five most recent reviews for each game posted as of Jan. 2012). Although we looked at games released in 2009, some user reviews were posted later.

Table 1. Number of Games in the Dataset by Platform

Platform	Number of Games	(%)
PS3	57	25.8
Xbox360	64	29.0
Wii	35	15.8
DS	37	16.7
PSP	28	12.7
Total	221	100.0

4. FINDINGS

Our goal was to determine a set of games for which we could compare their reviews. We wanted to identify groups of games with significant differences in critical reception between the US and Japan in order to then examine their reviews for insights that might explain the discrepancies. If a game was well received in the US but fared poorly in Japan, what were the reviewers raving (or ranting) about? Do these differences in scores (and reviews) reveal differences in opinion (e.g. some liked the camera controls while others didn’t), priorities (e.g. a game with poor controls is “punished” more severely), or something else (e.g. brand loyalties)? We were also unsure of the role hardware platforms might play. For example, the Xbox console performed quite poorly in Japan in terms of sales and “[a]nd discerning Japanese consumers hold foreign products to a higher standard than domestic competitors” [14]. Perhaps an Xbox360 game might be considered more harshly in Japan than its equivalent on the Playstation3?

We now present our findings. First, we will discuss the results of our statistical analysis of ratings scores. Next, we present the findings from our lexical analysis of game review texts. In each section we will complement and contextualize these findings with observations from our close reading of selected player reviews.

4.1 Statistical Analysis: Rating Scores

To start our analyses, we looked at the mean and the standard deviation (stdev) of the rating scores (Table 2 below). Note that scores are out of 100 points (for the four categories/sites). From the mean values, Japanese users (Game World) are the harshest critics (66.43) while Gamespot Users are the most lenient (77.90).

Table 2. Mean and Stdev of the Scores by Category

Category	Mean	Stdev
Metacritic	73.67	13.33
Gamespot Users	77.90	9.93
Famitsu (ファミ通)	72.57	9.41
Game World (ゲーム/セカイ)	66.43	12.13

¹ game.minpos.com

² www.metacritic.com/game

4.1.1 Inter- and Intra-culture Correlation

Next we calculated correlation coefficient (Pearson’s r) between the categories to see the inter- as well as intra-culture correlations (i.e., US vs. Japan, and critics vs. users). Table 3 shows the results. There is a strong (positive) correlation ($r > 0.5$) for all four pairings we examined. The strongest correlations occur within cultures (*A* for the US, and *B* for Japan), indicating at least a broad level of agreement between critics and users in the same culture. The correlation between the players (*D*) is distinctly lower than the rest suggesting there may be differences in appreciation of the same games between US and Japanese players. However, a weaker correlation is still a correlation – meaning that differences in appreciation may be subtle and not easily discernible.

Table 3. Correlation of the Scores Between Category

Review Score Pairing	Correlation
<i>A</i> : Metacritic vs. Gamespot Users	0.821
<i>B</i> : Famitsu vs. Game World	0.826
<i>C</i> : Metacritic vs. Famitsu	0.692
<i>D</i> : Gamespot Users vs. Game World	0.578

There are other explanations for the high intra-cultural correlation we observed. Generally speaking, user reviews are written after the appearance of reviews by critics. When users are writing their own reviews, they are often aware of the score of the “official” review. Livingston et al. found that players rated (scored) games significantly lower when they had read negative reviews prior to playing compared to players who had read a positive text [22]. Although their study focused on short-term effects and was designed “to eliminate influencing effects (e.g., anchoring)” [22] and other confounding factors, their findings may apply over a more extended period of time, perhaps more so for games a future amateur reviewer is looking forward to or has been doing research on prior to purchase. Another explanation for the high positive correlation is to consider an anchoring effect – a cognitive bias in which judgments are strongly influenced by an initial value [9]. Consider the “widget” for rating a game on Gamespot (Figure 1): it prominently displays the “official” score, Metacritic score and the average user score. Game World’s site is similar: users can submit scores and reviews from a pop-up page – with existing scores visible beneath. The bias due to an anchoring effect may result in users assigning higher scores when high numbers are shown (vice versa for low scores).



Figure 1 : Web Widget for Rating Games (Gamespot)

4.1.2 US vs. Japanese Player Rating Scores

We then examined the scores provided by users, both Gamespot and Game World, more closely. We wanted to see if there was

any difference between the users of the two cultures with respect to preferences on hardware platform. Table 4 shows the average score from each group of users broken out by platform. To get a sense of the differences between platforms, for each group and for each platform, we calculated how much the platform’s average rating was above (+) or below (-) the average of the rest of the platforms (combined) in the respective group.

Table 4. Average Rating Score by Platform

	Gamespot Users		Game World	
		(+/- mean)		(+/- mean)
PS3	79.44	+1.54	68.35	+1.92
X360	77.92	+0.02	67.03	+0.60
Wii	73.89	-4.01	62.74	-3.69
DS	78.41	+0.51	67.27	+0.84
PSP	79.07	+1.17	64.64	-1.79
All Games	77.90		66.43	

We did not find evidence of cultural biases against (or towards) specific platforms. It does not seem that Japanese players are harsher when it comes to rating games on Microsoft’s Xbox360 console (+0.02 vs. +0.60). However, there is a difference when it comes to games for Sony’s PSP. Game World users were more critical than Gamespot Users who were, relative to their respective mean, more lenient (-1.79 vs. +1.17). However the differences were not significant considering their standard deviations (9.93 and 12.13, shown in Table 2).

The differences in average ratings between platforms are explained by the quality of the games for them. We note the (comparatively) poor critical performance of games for Nintendo’s Wii.³ It has been noted that the Wii has many low-quality budget games (sometimes called *shovelware* [1]) that were “rushed out quickly to capitalize on the system’s popularity” [18, pg 72]. Further examination of the games is required to verify this.

4.1.3 Yoge

While there weren’t that many inter-cultural differences based on platform, we wondered if the same could be said when considering the “cultural origin” of each game. As mentioned earlier, *yoge* (洋ゲー) is a Japanese term used to refer to foreign (Western) games. The term is often used pejoratively. Would this be reflected in differences in review scores? We categorized all 221 games as “yoge” or “not-yoge. This categorization was done from a Japanese player’s perspective—we asked ourselves if they could consider the game as yoge or not. Since the term is used loosely, we used several criteria including: when was it released in Japan, was it developed by a Japanese company, does the game feature characters or IP strongly associated with Japanese popular culture (e.g. manga, anime), and is the game part of a franchise traditionally considered as originating in Japan. As a rule of thumb, we considered a game as not-yoge if it distinctly met any of those

³ The difference between the mean of the platform (Wii) and the mean of all other platforms (PS3+X360+DS+PSP) was statistically significant ($P < 0.05$) for both groups. The differences of all other platforms, for both groups, were not statistically significant.

criteria. This wasn't always easy: we categorized *Punch-Out!!* (Wii) as not-yoge despite being developed by a Canadian company because of how closely they worked with Nintendo during its development [25]. Similarly, we considered *Just Dance Wii* as "Not-Yoge" because, although it's a "Western" franchise developed and published by Ubisoft, the Japanese release was published by Nintendo, heavily localized, and featured a significant roster of Japanese popular music songs (J-Pop).

Table 5. Yoge vs. Non-yoge Rating Scores (by Game World)

	# of Games	Avg. rating
Yoge	99	69.40
Non-Yoge	122	64.02

Based on our interpretation of the term, we expected Japanese amateur reviewers would score yoge games more harshly. It was the opposite: Yoge games were rated higher than Japanese games. The difference between the means (69.40 for yoge and 64.02 for non-yoge) was statistically significant ($P < 0.05$). We speculated that this might be due largely to self-selection: users who wrote reviews on yoge games may be predisposed to view them positively regardless of the games' quality.

To look for evidence to support (or contradict) our speculation, we read the reviews in which the word "yoge" appeared. Among those reviews, some were for a yoge but others were not. Some mentions were a comment on the game being reviewed, while others were general statements on yoge. We were surprised to find many statements with negative sentiment on yoge. Here are some examples: "This game freezes very often. I wasn't so surprised because I am used to yoge, but...", "This game is not for those who don't like yoge or those who cannot deal with grotesque art." and "Texts are in larger fonts, unlike typical yoge." There were indeed positive comments as well, such as "This game is rough in terms of completeness but has a big content" and "Great graphics (as with many yoge)". But in the reviews in our dataset (Game World), there were more negative than positive comments on yoge. That means that Japanese users are not (comparatively) raving about yoge as the rating scores suggested.

Then why were yoge's rating scores high? One reason might be lower expectations. One example: "For a yoge, faces of the characters are not scary or grotesque – they are actually likable." Although the comment is positive, it betrays a negative preconception of foreign games: they have ugly characters by Japanese taste.

4.1.4 Best and Worst Games

Having analyzed the user rating scores from the perspectives of platforms and cultural origins, we looked at the distribution of the scores over all games to see the overall cultural preferences - which games received notably high or low rating scores in each culture. In addition to games which were highly liked or disliked by both cultures, we paid special attention to games which were liked by one culture but disliked by the other. To that end, we sorted the games based on the rating score within each culture, and categorized as 'Best' and 'Worst' those games which were more than +/- 1 standard deviation away from the mean. By using the cut-off, there were 28 Best games for Gamespot and 26 Best games for Game World, and 25 Worst games for Gamespot and 35 Worst games for Game World.

As we noted earlier, Wii games were significantly worse rated in both cultures. A closer look at the games in the 'Worst' lists was

somewhat surprising since of the 12 Wii games (6 in Gamespot and 11 in Game World), only one or two could be called "shovelware". Notable critical failures include *Castlevania Judgment* and *Major Minor's Majestic March*.

For the "Best" list, in the case of GameWorld, we noted that more than half of the games are yoge. Notable on the list are *NBA2K10* (PS3), *DiRT 2* (PS3), *Trials HD* (X360), *Tiger Woods PGA Tour 09 All-Play* (Wii), *Fallout 3* (X360 and PS3), *Uncharted 2: Among Thieves* (PS3), *Grand Theft Auto: Chinatown Wars* (DS). The yoge games that made the list represent a reasonably diverse number of genres (sports, RPG, action, driving) and "size" – not all are AAA blockbuster titles.

4.1.5 User Rating Agreement

To further analyze the difference in preferences between the users of the two cultures, we examined their rating agreement. Table 6 shows a contingency table indicating the agreement/disagreement between the ratings by Gamespot and Game World users. In the table, games are binned/partitioned based on their review scores and how they compared to the mean. From left to right (columns) we counted how many games' review scores from GameWorld had: very poor ratings (one standard deviation below the mean), poor ratings (below the mean, but no worse than one standard deviation), good ratings (above the mean, but not better than one standard deviation), very good ratings (more than one standard deviation above the mean). We then did the same for the rows according to the user reviews on Gamespot. In this way, the top-right corner indicates that there were 17 games whose review scores were more than one standard deviation lower than the GameWorld mean score (left column) and more than one standard deviation below the mean Gamespot user score (top row). These 17 games were the lowest rated in both cultures. The table's diagonal (values of 17, 32, 64, and 13) represents those games for which there was the highest degree of agreement in both cultures. Since there were no games for which there were serious inter-cultural disagreements (zeros in the bottom left and top-right corners), we decided to look at the next closest level of disagreement. In Table 6, these are the shaded cells.

Table 6. Cross Tabulation of Games by User Review Scores

		GameWorld			
		Very Poor ^a	Poor ^b	Good ^c	Very Good ^d
Gamespot Users	Very Poor ^a	17	3	5	0
	Poor ^b	11	32	21	3
	Good ^c	7	20	64	10
	Very Good ^d	0	2	13	13
		a: Review was more than one standard dev. below the mean b: Review was below the mean, but within one standard dev. of it c: Review was above the mean, but within one standard dev. of it d: Review score was more than one standard dev. above the mean			

Our close reading of game reviews (both in Japanese as well as in English) highlighted a variety of issues. Players in both groups tended to praise (or condemn) similar kinds of issues in all games: graphics, controls, and so on. There were, however, two kinds of issues that seemed to occur with greater frequency in Japanese (negative) reviews than they did in US (negative) reviews: technical issues and apparent problems with translation or localiza-

tion. We saw Japanese players often reporting problems with bugs and having connectivity issues (online lag). This is something we did not see as often in the case of US user reviews. It is possible that US player’s expectations are different with regards to bugs and problems in games. Bugs are sometimes “welcomed” by some player communities for several reasons including players simply being excited about finding them [2]. Humor is also important with several online sites dedicated to posting videos that showcase bugs and glitches in popular games. Not all bugs are tolerated equally, but generally the expectation seems to be that the developer will issue patches and fixes to solve problems. We note that this probably is not an issue of the Japanese versions of games being more “buggy” than their Western releases, since we found evidence of similar complaints in reviews from both sites. It may be that Japanese gamers just are not as tolerant of buggy software and more punishing in their reviews as a result of this.

Complaints about translation and localization were also frequently brought up in Japanese (negative) reviews. Consider *Call of Duty: Modern Warfare 2*, several Game World reviews complained about errors in the translation (to Japanese) and how they weren’t impressed with the quality of the Japanese voice actors. Similarly, for *NBA2K10*, reviewers noted problems with typos and inconsistencies in the story. We note that sometimes these issues (possibly bad translations) may be conflated with other aspects of a game. When a reviewer is commenting on their dissatisfaction with a game’s narrative, is it perhaps the case that there were translation problems that rendered the story incoherent? It is plausible that the poor reception of some games, despite critical acclaim in the West, is due to localization issues. Another possibility is that, as we believe might be the case for technical problems and bugs, Japanese players are more sensitive to these kinds of issues.

4.2 Lexical Analysis: Word Frequency

So far we have analyzed our dataset statistically. Now we will present results from our lexical analysis. Here we analyze the words used in the user reviews to identify general aspects of games which the users like or dislike in each culture. By looking at words from all reviews in our dataset, especially the frequency of words, our aim was to extract a broad range of game aspects that users care about. This is based on the assumption that people will talk more about things they feel strongly about, positively or negatively. So, the words with high frequency would indicate the things/aspects which they are sensitive to. Then, by comparing the aspects extracted from the two cultures, we can discover the differences in their game appreciation and preferences.

In our word extraction, we focused on nouns as a part-of-speech (POS). We chose nouns because words that express game aspects (e.g. graphics, gameplay, voice acting) are mostly nouns (and noun phrases). In order to associate game aspects with preferences (i.e., which aspects were liked/disliked) we divided the reviews from each site (Gamespot and Game World) into two sets: positive and negative. For the Gamespot reviews we determined a review was positive if it gave a rating score of 8.0 or above, or negative otherwise. We established this cut-off because Gamespot considers games scored at 8 or above as “Great”: a clearly positive review. Since GameWorld user reviews have separate good and bad sections (see Figure 2), we used these sections from all 4,729 reviews to create the positive and negative sets.⁴

⁴ In the 1,045 Gamespot reviews, 631 were positive and 414 were negative.

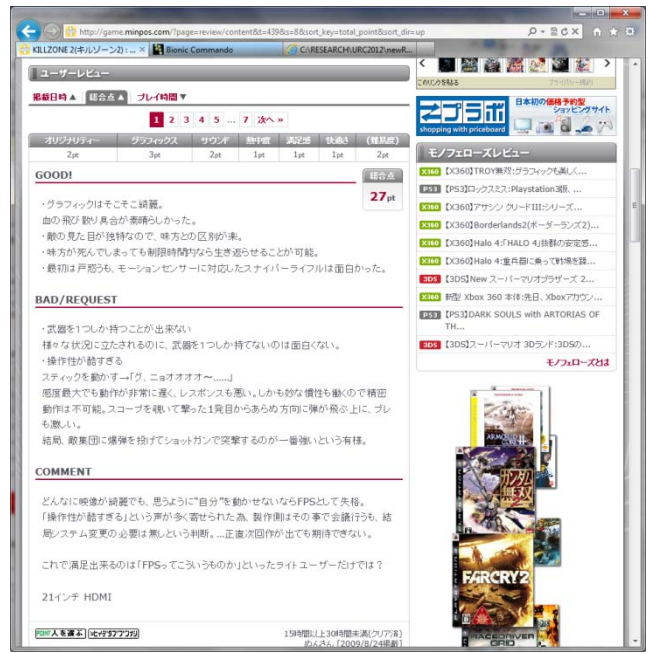


Figure 2. Example User Review from Game World

For each of the four sets we then automatically extracted nouns and noun phrases⁵ and examined the frequencies for each word. There were a total of 12,871 and 9,545 nouns/nouns phrases in the positive and negative sets respectively for Gamespot, and 6,738 and 10,676 nouns/nouns phrases in the positive and negative sets respectively for Game World. Table 7 shows examples of the words and phrases which appeared notably frequently in either positive/negative set (or both) for each culture.⁶

Table 7. Words Frequently Used in Positive/Negative Reviews

Gamespot	
Pos	character, story, graphic, player, weapon, world, battle, boss, race, music, action, sound, multiplayer, ability, challenge, team, quest, skill, style, puzzle, speed, chance, replay value, blood, combo,
Neg	time, enemy, level, gameplay, attack, hour, problem, control, combat, environment, score, screen, variety, stage, match, voice, button, lack, idea, minute, map, visual, option, camera, dialogue, machine,
Game World	
Pos	graphic, system, music, series, action, world, sound/tone, atmosphere, production, expression, cooperation, race, movie, tension, background, art, effect, realism, completeness, moving, RPG,
Neg	enemy, character, attack, story, item, operation, control, hour, time, difficulty, stage, level, movement, screen, save, problem, specification, mission, bug, display, viewpoint, camera, freeze, lock-on, lag,

⁵ We used a Natural Language Processing (NLP) tool called *Stanford CoreNLP* (<http://nlp.stanford.edu/software/corenlp.shtml>).

⁶ Words in this table had a very high frequency count in either set, or the ratio between the frequencies in the two sets (i.e., Good/Bad, Bad/Good) was notably lopsided. Words in each cell in the table are roughly in descending order of frequency.

There are many observations we can make from the table. One observation is that both cultures have graphics/art and music/sound on the positive side, suggesting that games with good audio-visuals are well received by both cultures. Another is that words such as time, enemy, level, control, screen and their related words are on the negative side for both cultures, suggesting that games which take an unnecessarily long time, or have bad game design or interface, or difficult controls, are commonly disliked by both cultures.

There are also words which are on one side in one culture but on the opposite side in the other culture. For example, the word “story” is on the positive side in Gamespot but on the negative side in Game World. To find out the reason, we randomly selected several reviews and found sentences in which “story” appeared. Gamespot users were indeed often referring to the story of the game (being good) along with other good aspects in positive reviews (e.g. “it draws you in with a fantastic story and a set of great characters.”), but the opposite was true in Game World (e.g. “Story is rough”, “The game is too much story-based”). Ngai’s findings suggest that Japanese players value character and story development in games quite highly [23]. Since Japanese players place more emphasis on story, their overall appreciation of a game is strongly dependent on the quality of its narrative. A bad narrative (including one poorly translated) would have a higher impact on the overall evaluation of a game.

There are several words which are on one side in one culture but on neither side in the other. Bug, lag and freeze are often mentioned in the negative reviews in Japanese negative reviews but hardly at all in the US. This result seems to support our earlier speculation that Japanese users may be less tolerant of games with technical issues. We also note the appearance of “replay value” on the positive side in Gamespot. Western reviewers often comment on the “replayability” of a game: a combination of how compelling a game is to play it again (and again) and how many additional things there are to do in a game once the main storyline has been completed. More is better. This concept doesn’t seem to figure in Japanese user reviews, or at least it was not explicitly spelled out or articulated in their review texts.

The overall worth of a game is evident in Japanese reviews in a subtly different way. The word “completeness” appears frequently in positive GameWorld reviews and is used to refer to whether or not everything the game should have is there, together with how well everything has been executed. The use of the term is similar to “polish” – whether or not a high level of quality and execution is present in all aspects of a game. The word “movement” often appears on the negative side in GameWorld reviews because reviewers complain about un-natural movement and animation (i.e. not realistic), and thus the game is not well made, rough, incomplete, and so on.

Another interesting example is that Gamespot has the word “score” on the negative side, but the word did not appear on either side in Game World – which suggests that Japanese users are not as sensitive to the in-game score as US users.

5. Game Reviews: Further Observations and Future Directions

While online communities, both videogame related and not, have been a frequent subject of study in recent years, this hasn’t been the case for reviewing communities [11]. It is perhaps for this reason that we were often surprised by what we read. Our goal in this section is to provide additional context for our findings as well as describe some unexpected things we noticed. We note that

these observations should not be considered as conclusive findings – rather they outline areas of interest that should be examined more deeply and systematically in the future.

5.1 Review Process

In education, the term scaffolding is used to refer to support for learning “that communicates process, coaches, and elicits articulation” [13]. We wondered how the design of the websites we studied may communicate what a review is or should be and how they may coach their reviewers into what to write.

Writing a review on Gamespot requires logging in, finding a game and reaching a “Reader Review” page. At the top of the page are several guidelines including a minimum length (100 words) and an admonition that opinions should be supported by facts: “Merely stating ‘this map pack sucks’ or ‘this horse armor is the best’ is not enough” [10]. The review is divided into three sections. The first, Scoring, includes a slider for setting the score and three pull-downs (difficulty, time played, “classification”) with multiple options to select from. The third pull-down’s options are short phrases such as “Grows on you” and “Rent it first” meant to represent an overall view of the game. The second section, the “Review Deck”, is a textbox with a 120-character limit. This text is what appears in the reviews index and should summarize what is written in the 3rd section, the Review Text (also a textbox). Finally, at the bottom, there are two buttons: spell check and submit. The order of these elements is important. By starting with the score, reviewers are more likely to focus on how their soon-to-be-written review reflects the score and classification they’ve already chosen. In this case the review should justify the score already given, rather than have the score summarize the review.

For user reviews on GameWorld, the emphasis seems reversed. Here, the submission page is longer and with more sections. Beneath some guidelines similar to those on Gamespot, readers must fill out some personal details (e.g. name, email). After this there are three textboxes, each requiring a minimum of 50 characters. In order, these are “Good”, “Bad/Request” and “Comment”. Following the textboxes and further guidelines, there are seven pull-downs for assigning scores. Scores can range from 0 to 5, although reviewers are discouraged from both extremes of the scale – these should be used in special cases only. The seven aspects are “Originality”, “Graphics”, “Sound”, “Degree of Enthusiasm”, “Sense of Satisfaction”, “Sense of Comfort”, and “Difficulty Level”. Finally, reviewers can indicate how long they’ve played the game, if they’ve completed it, and select up to three tags from a group of 24 (e.g. “Good Combat”, “For Beginners”, “Recommended to play prequels”). Game World’s mandatory “Bad/Request” section requires that its users describe some negative aspects of the game (in at least 50 characters). Also, with the scoring at the end, it is likely that the scores reflect what was described in the review rather than have the score lead the writing. Also, in terms of the final score, some games may be comparatively punished due to weaknesses in a particular area – players may love a game despite its poor graphics, but they have no way to “sweep that under the rug”. This might explain the apparent harshness in Game World user reviews (Table 2). While Game World’s user review page encourages reflecting on various aspects of a game, it may discourage assuming a broader view.

5.2 The Practice of Game Reviewing

Given similar qualitative descriptions, reviewers may award very different numerical scores [20]. We found examples of this. We also found instances of apparent internal inconsistency. These were cases in which a review’s text seemed to reflect one evalua-

tion, but the numerical score contradicted it. In the summary shown before reading the full text (see Figure 3), the reviewer has described the game as “not great”, yet assigns a score that is interpreted by Gamespot’s scale as “great”. This is not too surprising; since Gamespot reviewers seem quite lenient when it comes to scores. Further research is needed to understand why this happens.

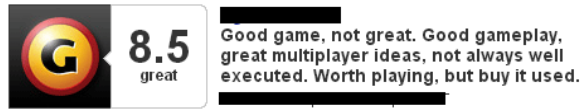


Figure 3. Example of Inconsistent Review (anonymized)

Inconsistent reviews (i.e. rating scores do not follow the text) can be problematic, especially in determining a review as positive or negative. However, the analysis method we employed was effective in circumventing this issue. By taking a large body of reviews, we can assume that inconsistent reviews are not as common and won’t skew the results. Also by manually inspecting the text of the reviews which induced a significant or counter-intuitive result, we can detect and filter out erroneous reviews.

We think it is possible that, due to game reviews forming a central part of gamer culture and identity [26], players have a strong shared understanding of what different scores mean and how they should be considered. Professionally written reviews are a part of the ecosystem in which they are writing their own reviews – so non-professional reviewers may be modeling their own reviews on what they are reading on the site. This could be seen in some of the general reviewing conventions some authors employ without prompting from the sites’ interface or guidelines. Consider the following excerpt from Gamespot (edited for length):

*“Story: 5/5 (Good storytelling)
 New features: 5/5 (Noise change is great)
 Graphics: 4/5 (Good but not good enough)
 Music 5/5 (Great music score)
 Battle system: 4/5 (Most battle cards remain the same)
 Multiplayer: 5/5 (Many Multiplayer features)
 Gameplay: 5/5 (Almost like the battlenetwork series)
 Overall: 4/5 (A game for [GameCharacter] fans)”*

Here the author has broken down the review into areas/sections including a “final score” meant to reflect the overall rating of a game. Reviewing a game by considering its “parts” is common in the gaming press. However, the breakdown offered by the author above is not typical, nor is it the one used on the site on Gamespot. The breakdown above (which came after a paragraph of descriptive text) highlights two aspects the author felt were particularly relevant: “new features” and “battle system”. The former signals that the game is part of a series and alerts players of earlier games whether they should be interested while the latter (“battle system”) places additional emphasis on one of the games core gameplay features such that it should be considered separately from regular “gameplay”. It was also common for both US and Japanese reviewers to comment on their reviewing experience with comments like “this is my 3rd review for my newest psp game” or “Ok, as I do in every one of reviews...”. Player reviews are quite conversational, offering direct advice to the reader and inviting further participation. Comments like “stay tuned for my next review” illustrate the assumption that readers will engage in a

conversation with the reviewer (such as following, tracking an author, and rating the review). Furthermore, there is also often a reference to a broader community, a reviewer might note how “I have to agree with everyone here” or “I don’t seem to understand the reviews of this game. I’m sorry, but it seems that no one has reviewed it for what it is...”

What we might be seeing here is that, contrary to other sites where “amateurs write to scratch an itch” [11], we have writers for whom reviewing is primarily about establishing an identity and credibility as a gamer. Who you are is determined by what you’ve reviewed (and thus played), and the quality of your review. This might explain why so many player reviews dedicate significant space to describe the author’s experience and opinion of other games. Consider this excerpt (edited for grammar and anonymity):

“I’ve been a huge fan of [Game2009] ever since I played [Game1996] on the PSOne. I was quite excited for [Game2009] but in 2009 I didn’t have a PS3. [...] I’ve completed it on the 360, but unfortunately I didn’t feel like reviewing it. In 2011 [...] I rented it on PS3 and played mostly the online mode. [...] But recently I bought the PSP copy and had even more fun.”

We also found US user-submitted reviews written prior to a game’s US release. The authors presumably imported the game, played it, and then posted their review. Rather than treat the review as a “scoop” (i.e. we beat Gamespot), they seemed to consider their work as a service for the hard-core fans anxiously waiting to hear about the game. Along the way, they establish a reputation within the community.

This notion of credibility is cross-cultural. The 3rd section of the reviews on Game World (“Comment”) is for reviewers to provide a summary. It is also often used by reviewers to refer to their experience with a game series as well as their “play environment”. For example one reviewer notes how they “Have played all the prequels of this game” and that they “played on a CRT-based TV.” Another reviewer explains that they played using a “24 inch HDMI LCD monitor connected with 5.1 headphones” and that it “took around 12 hours to clear all levels one time through.” This information lends credibility to the reviewer while also strengthening their identity in the community.

6. CONCLUSIONS

We have presented our findings on cultural differences in the appreciation of games based on user reviews of games written in Japan and in the US. We obtained several findings, including that users in neither culture have bias for or against particular platforms and that preferences of the users in the two cultures are generally similar. By reading the reviews of the games which received very high/low ratings, or games for which the ratings disagreed between the two cultures, we found that Japanese users are sensitive to bugs and problems related to localization. Also by examining words which appeared frequently in positive/negative reviews, we found that American players tend to emphasize replay value in games. We also obtained a finding on yoge that was contrary to our intuitions. Based on the rating scores, Japanese players seemed to favor yoge. However, closer inspection of the reviews revealed inherent negative perceptions on yoge.

Our study does have limitations. One significant limitation might come from the dataset. Analyzing responses to the same games allows for comparisons, but it is possible that the most interesting differences in game appreciation are conveyed precisely in those games that are not exported (or imported). However, as the game

industry has globalized, companies are creating games with an international audience in mind. So, the number of games that are unique to one culture has been decreasing and will continue to be so. Also, it is important to consider the scope of non-exported games – they might represent examples of small niches, rather than broad acceptance even within their culture.

For future work, we plan to collect newer reviews and do a temporal analysis to examine the changes in the cultural preferences. The goal of the research would be to investigate the effect of game globalization. Another topic we are planning to focus on is localization. As for analysis methods, we are planning to develop a computer application which predicts rating scores for games. The goal is not so much to build an application per se, rather to build a model which simulates human players' rating scores and learn the parameters involved in the model.

7. ACKNOWLEDGMENTS

We would like to thank Matt Parker and Zach Luttmner for early data collection, Kosuke Yabumoto for helping process Japanese reviews, and Sigrid Norris for comments on an early draft.

8. REFERENCES

- [1] Apperley, T., 2012. Nintendo Wii. In *Encyclopedia of Video Games: The Culture, Technology, and Art of Gaming: Volume 1*, M.J.P. Wolf Ed. Greenwood, Santa Barbara, CA, 462-464.
- [2] Bainbridge, W.A. and Bainbridge, W.S., 2007. Creative Uses of Software Errors: Glitches and Cheats. *Social Science Computer Review* 25, 1, 61-77.
- [3] Bond, M. and Beale, R., 2009. What Makes a Good Game?: Using Reviews to Inform Design. In *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology* British Computer Society, Swinton, UK, 418-422.
- [4] Consalvo, M., 2006. Console video games and global corporations: Creating a hybrid culture. *New Media & Society* 8, 1, 117-137.
- [5] Consalvo, M., 2007. Visiting the Floating World: Tracing a Cultural History of Games Through Japan and America. In *Situated Play, Proceedings of DiGRA 2007 Conference* DiGRA, Tokyo, Japan.
- [6] Cook, G., 2009. "Methods to Market Mario: An Analysis of American and Japanese Preference for Control in Video Games." from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1347411.
- [7] Edge Staff, 2010. "Vanquish Preview." Edge, Retrieved September 3, 2012, from <http://www.edge-online.com/features/vanquish-preview?page=show>.
- [8] Farrell, S., 2012. Metacritic. *The Charleston Advisor* 13, 3, 26-31.
- [9] Furnham, A. and Chu Boo, H., 2011. A Literature Review of the Anchoring Effect. *The Journal of Socio-Economics* 40, 35-42.
- [10] Gamespot, 2012. "Reader Review." Gamespot, Retrieved 12 Dec, 2012.
- [11] Gilbert, E. and Karahalios, K., 2010. Understanding Deja Reviewers. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work (CSCW '10)* ACM, New York, NY, 225-228.
- [12] Gordon, J., 2009. "Peter Moore: 'Breaking Japan is Hard'." Now Gamer, Retrieved September 3, 2012, from http://www.nowgamer.com/news/914882/peter_moore_breaking_japan_is_hard.html.
- [13] Guzdial, M., 1995. Software-realized scaffolding to facilitate programming for science learning. *Interactive Learning Environments* 4, 1, 1-44.
- [14] Hayes, J., 2005. Reconceptualizing the Xbox Platform. *Design Management Review* 16, 4, 10-17.
- [15] Huber, W., 2008. Notes on aesthetics on Japanese videogames. In *Art and Videogames*, A. Clarke and G. Mitchell Eds. Intellect Books, London.
- [16] Iwabuchi, K., 2002. "Soft" Nationalism and Narcissism: Japanese Popular Culture Goes Global. *Asian Studies Review* 26, 4, 447-469.
- [17] James, K., 2010. Mario vs. the Lich King: How Culture Affects American Consumers' Preferences for American or Japanese Video Games. In *Dietrich College of Humanities and Social Sciences* Carnegie Mellon University, Pittsburgh.
- [18] Jones, S.E. and Thiruvathukal, G.K., 2012. *Codename Revolution: The Nintendo Wii Platform*. MIT Press, Cambridge, MA.
- [19] Koyama, Y., 2007. Survey of the adjourned sale rate of the Japanese home video game industry. In *Situated Play, Proceedings of DiGRA 2007 Conference* DiGRA, Tokyo, Japan.
- [20] Kushal, D., Lawrence, S., and Pennock, D.M., 2003. Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews. In *Proceedings of the 12th International Conference on World Wide Web (WWW '03)* ACM, Budapest, Hungary, 519-528.
- [21] Livingston, I.J., Nacke, L., and Mandryk, R.L., 2010. Critic-Proofing: Robust Validation Through Data-Mining. In *Proceedings of Fun and Games 2010 Workshop at Playability and Player Experience Conference* (Leuven, Belgium 2010), NHTV Expertise Series, 81-94.
- [22] Livingston, I.J., Nacke, L.E., and Mandryk, R.L., 2011. The Impact of Negative Game Reviews and User Comments on Player Experience. In *Proceedings of 2011 ACM SIGGRAPH Symposium on Video Games, Sandbox '11* (Vancouver, Canada 2011), 25-29.
- [23] Ngai, A.C.Y., 2005. Cultural Influences on Videogames: Player's Preferences in Narrative and Game-Play. Master's thesis. University of Waterloo, Waterloo, Ontario. <http://hdl.handle.net/10012/770>.
- [24] Staff, 2012. Japan's Game Media King. In *Retro Imagine* Publishing, UK.
- [25] Totilo, S., 2009. "Punch-Out Devs Talk Graphics, Difficulty, Nixed Princess Peach Idea." Kotaku.com, Retrieved December 6, 2012, from <http://kotaku.com/5320198/punch-out-devs-talk-graphics-difficulty-nixed-princess-peach-idea>.
- [26] Zagal, J.P., Ladd, A., and Johnson, T., 2009. Characterizing and Understanding Game Reviews. In *Proceedings of the International Conference on the Foundations of Digital Games* ACM, New York, NY, 215-222.
- [27] Zagal, J.P. and Tomuro, N., 2010. The Aesthetics of Gameplay: A Lexical Approach. In *Proceedings of the 14th International Academic Mindtrek Conference, Tampere, Finland*, A. Lugmayr, et al. Eds. ACM, 9-16.