Sentiment Analysis of Souls-like Role-Playing Video Game Reviews

Tibor Guzsvinecz

Department of Information Technology and its Applications, University of Pannonia, Gasparich Márk utca 18/A, 8900 Zalaegerszeg, Hungary
guzsvinecz.tibor@zek.uni-pannon.hu

Abstract: The emotions and experiences of players are complex as they can vary due to many factors, such as the difficulty of a video game. However, after finishing video games, most players write reviews of them in which their experiences are described. To understand these better regarding challenging video games, sentiment and textual analyses were conducted on Steam reviews. 993932 reviews were scraped from its website and were imported into R. Using several packages available in R (syuzhet, dplyr, tidytext, ggpubr) and the NRC Emotion Lexicon, the data were investigated. Reviews were grouped into negative and positive ones. According to the results, the following can be concluded regarding player experience in challenging video games: 1) Significant differences exist in case of all emotions and sentiments between the two groups; 2) The emotional valence among both groups significantly differs as well: of negative reviews, it is below zero, and of positive ones is above it; 3) In case of positive reviews, the commonly used words better detail player experiences than in case of negative ones; 4) Regarding the number of votes on and words in both negative and positive reviews, no correlation can be found between them. However, negative reviews are more likely to be agreed with.

Keywords: cognitive process; data analysis; emotion; player experience; sentiment; souls-like; video game

1 Introduction

Emotion can be defined as a complex reaction pattern since it involves physiological, experiential, and behavioral elements, by which an individual attempts to deal with a personally significant event or matter. The significance of events determines the specific quality of the emotion. Among others, this quality can be anger, fear, joy, and surprise.

Research regarding emotions dates back to the 19th Century [1] and is still being studied today. Some concluded that cognition is a part of emotions, while others claim that it is separate from them. However, newer studies indicate that emotions are related to cognition: according to the study of LeDoux and Brown, emotions are cognitive states resulting from information gathering [2]. The study of Larue et al. shows that emotions play an important role in cognition [3].

This is quite important because in today's world almost everything can be experienced. Consequently, certain emotions can be evoked and video games are no exception [4]. The study of Scoresby and Shelton shows that a video game can create motivation since it emotionally links the players to the content [5]. The player's feeling of presence and immersion can also be affected [6, 7]. Consequently, when someone is playing, a cognitive process occurs. This happens because players perform actions and then they evaluate the outcomes. Players refine their behaviors by interpreting and reflecting on the feedback that the game gives them. Behavior patterns can also emerge [8]. Therefore, some sort of cognitive process occurs for this evaluation [9]. Processing game data perceptually is also necessary. This whole process is called player experience (PX) [10]. Naturally, visual and auditory output is crucial for this process since players have to be informed of the changes that happen in the game environment.

This is where the field of Cognitive InfoCommunications (CogInfoCom) comes into the picture as it investigates the interaction between the human and the machine [11-14]. It aims to improve, return or even create new cognitive abilities through models based on the ICT engineering tools [15-20]. Consequently, emotions and user experience (UX) are also investigated in it [21-27], and the latter is extremely similar to PX.

Similarly to the case of UX, PX is also quite complex: it can be shaped by several factors, such as difficulty settings, game mechanics, narrative, soundtrack, or even by reading game reviews [28-33]. PX can also be impacted by emotional triggers such in-game loss, character bonds, and personal memories [34]. Since players frequently share their experiences in their reviews, it is possible to understand PX to some extent [35]. The fact that hundreds of reviews are posted every day gives academics access to a lot of data [36]. The analysis of textual reviews can reveal important details about the players and games. Busurkina et al. state that the following topics can be observed in reviews: general experience, achievements, social interaction, social influence, accessories, visual/value, narrative, and bugs [37].

To understand the experience and emotions of players in case of challenging games, textual reviews of difficult video games on the Steam platform were scraped and analyzed. Analyzing the reviews on Steam is not a foreign concept as its API allows one to scrape them. Furthermore, several studies examined them in the past [37-42]. According to Kang et al., both the reviews and votes on them by other players on Steam are quite useful [43].

Therefore, this paper is structured as the following: the materials and methods are presented in Section 2, while the results are shown in Section 3. Discussion can be found in Section 4, and after that, conclusions are drawn.

2 Materials and Methods

Before the scraping commenced in April 2021, challenging video games had to be selected. After an empirical search, the so-called "Souls-like" role-playing video game genre was chosen for this and a previous study [42]. Games in this genre are generally praised for their unforgiving difficulty. Therefore, the inclusion criterion is met because these are games that challenge their players.

"Souls-like" games are often 3D role-playing games with a third-person camera. Since they vary in design, 2D side-scrolling versions of them exist as well. A few of them can be seen in Figure 1 as examples. However, what makes a video game "Souls-like"? First, the unforgiving difficulty and consequently, a unique checkpoint system. Imagine that the goal is to reach and defeat the final boss of the game. The path to it is difficult, as these types of games do not tell the player where to go. Naturally, there are environmental hazards and enemies on the way. These can easily vanquish the player's character with a few hits or missteps. If the player character is vanquished, it is returned to the last checkpoint. In this case, all enemies (except bosses) come back to life, and the collected currency is dropped on the spot. If the player's character does not pick up the dropped currency before being vanquished again, it is lost forever. The second common characteristic of a "Souls-like" game is the contextual/environmental storytelling. The narration of these games is usually cryptic, and the player has to search for the story in the environment. It is possible to finish these games without knowing their story.

Figure 1
Examples of "Souls-like" games: Dark Souls III (upper left), Death's Gambit (upper right), Hollow Knight (lower left), and Nioh: Complete Edition (lower right)

21 “Souls-like” games were chosen for the research using the user tag system on Steam. However, games had to be selected carefully as the tags are placed by users. Due to this, the “Souls-like” tag could be found on games that are not “Souls-like”, and even on video editing software. Games that were selected vary in design, but their genre remains the same. They can be considered new as well since their release dates are between 2011-2020. It should be noted that more of these games are still in development today, but early access games and those that are new, thus have few reviews, were not included in this study.

After the selection, all (993932) reviews at that time were scraped using the steam_reviews Python package, which is available under the MIT license [43]. The reviews were downloaded in a JSON format using this package. For the analysis, the JSON files were imported into the statistical program package R [44] with the help of the jsonlite package. The scraped reviews per video game can be found in Figure 2.

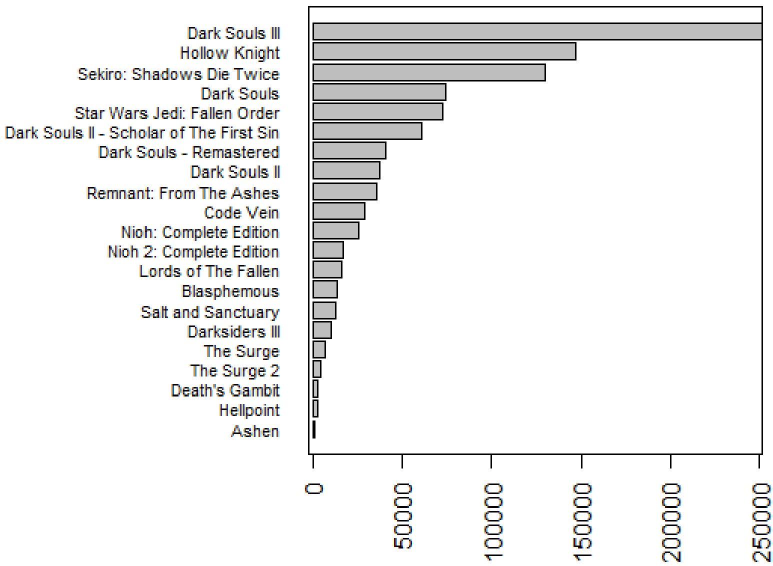

Figure 2

The number of scraped reviews per video game

2.1 Reviews on Steam

On Steam, each video game has a page, and the reviews can be found on the games’ subpage. Several fields of information are contained in a single review [45], but only the following were used during this research: the name of game, the language of the review, its textual part, whether the reviewer gave the game a thumbs up or a thumbs down, and how many players “agreed with” each review.

It should be noted that the Steam reviews do not have a rating system between a scale of 1-10 (or 1-5). This means that every review is either positive or negative. When reading reviews, this can be easily seen as it is symbolized with a thumbs up or a thumbs down on each one. Naturally, thumbs up means that the reviewer recommended the game, while in the case of thumbs down, the game is not recommended. It should be noted these are used as synonyms in this study. Thus, a “positive review” means “thumbs up” and “recommended game”; while “negative review” means “thumbs down” as well as “not recommended game”.

Out of the 993932 scraped reviews, 904005 (90.96%) are positive, and 89927 (9.04%) are negative. In this study, only English reviews were investigated, thus these numbers decreased to 377334 positive, and 41149 negative reviews. The percentage of positive English reviews is 90.16%, while of the negative English ones, it is 9.84%. The descriptive statistics of these reviews can be observed in Table 1.

Table 1
The number, and percentage of English reviews grouped by game and review type

Game	Positive review	Negative review
Ashen	515 (70.93%)	211 (29.07%)
Blasphemous	5339 (89.16%)	649 (10.84%)
Code Vein	11903 (90.89%)	1193 (9.11%)
Dark Souls	39864 (90.35%)	4254 (9.65%)
Dark Souls – Remastered	16996 (86.53%)	2645 (13.47%)
Dark Souls II	16285 (86.13%)	2622 (13.87%)
Dark Souls II – Scholar of The First Sin	19737 (83.57%)	3879 (16.43%)
Dark Souls III	88320 (92.93%)	6714 (7.07%)
Darksiders III	3204 (77.14%)	949 (22.86%)
Death’s Gambit	1170 (74.42%)	402 (25.58%)
Hellpoint	1184 (80.54%)	286 (19.56%)
Hollow Knight	58029 (96.98%)	1805 (3.02%)
Lords of the Fallen	3298 (54.03%)	2805 (45.97%)
Nioh: Complete Edition	6291 (83.07%)	1282 (16.93%)
Nioh 2: Complete Edition	4524 (89.60%)	525 (10.40%)
Remnant: From the Ashes	16729 (87.30%)	2432 (12.70%)
Salt and Sanctuary	4461 (89.99%)	496 (10.01%)
Sekiro: Shadows Die Twice	33795 (92.74%)	2644 (7.26%)
Star Wars Jedi: Fallen Order	41130 (90.93%)	4102 (9.07%)
The Surge	2720 (74.58%)	927 (25.42%)
The Surge 2	1840 (84.91%)	327 (15.09%)

2.2 Textual Analysis

To analyze the emotions in the textual parts of the reviews, a Natural Language Processing package called syuzhet was used in the statistical program package R [46]. While this package incorporates four sentiment lexicons, the NRC Emotion Lexicon was chosen since it is free for research purposes. It is also proven to be useful in the literature [47-49].

There are eight basic emotions and two sentiments in the NRC Emotion Lexicon. The emotions are anger, fear, anticipation, trust, surprise, sadness, joy, and disgust, while the sentiments can be positive or negative. The syuzhet package has a customizable function called get_nrc_sentiment with four arguments. Using this function the sentiments in the text can be assessed. Afterward, a data frame is created: each row corresponds to a sentence, while the columns represent the emotions [50]. The negative numbers from the respective column are converted and added to the values in the positive column. The number of sentiments per sentence can be found in the resulting matrix as can be observed in Table 2, although it is transposed to fit into the margins of the paper. Using this method, the emotions and sentiments were investigated as can be seen in subsection 3.1.

Table 2
Examples of the number of sentiments in sentences. The table is transposed compared to the data. The sentences were: (1) This game is very good. (2) A great Souls clone! Love this game! (3) Love this "free roaming" Dark Souls clone. Beautiful graphics and level design.

Emotion	Sentence 1	Sentence 2	Sentence 3
Anger	0	0	0
Anticipation	1	0	0
Disgust	0	0	0
Fear	0	0	0
Joy	1	1	2
Sadness	0	0	1
Surprise	1	0	0
Trust	1	0	1
Negative	0	0	0
Positive	1	1	3

The emotional valence in reviews was also investigated (subsection 3.2). However, since the reviews have a variable length, percentage-based chunks were used. With the help of the get_percentage_values function, the text was divided into an equal number of chunks, and the mean emotional valence was calculated for each. 50 chunks were selected for this study after empirical testing.

To calculate word frequencies in subsection 3.3, two packages were used: dplyr [51], and tidytext [52]. After the frequencies were found, the package called wordcloud was used to plot the most common words [53].

It should be noted that on Steam, it is possible for every user to “agree with” (in other words: “vote on”) the reviews of others. This means that besides voting, players can see and read the reviews of others. If they vote, it would mean that they agree with the reviews, naturally. The correlation between the length of the reviews and the number of votes on them was investigated. The results of this investigation can be found in subsection 3.4. This was assessed in detail since according to Kang et al. [41] the number of votes on reviews is important since these “voted on reviews” could mean the opinions of multiple players.

3 Results

Since four different analyses were conducted, this section is split into four subsections. The difference in emotions and sentiments was investigated in subsection 3.1, the difference between emotional valence during the narrative time was examined in subsection 3.2, the difference among frequently used words is analyzed in subsection 3.3, and lastly, the difference in agreeing on reviews is looked at in subsection 3.4.

3.1 Difference in Emotions and Sentiments

First, the sentiments were investigated in the reviews. Two groups were made from the data: positive, and negative reviews. Using the NRC Emotion Lexicon, the following percentage of the average valence of basic emotions can be found in the reviews (Figure 3):

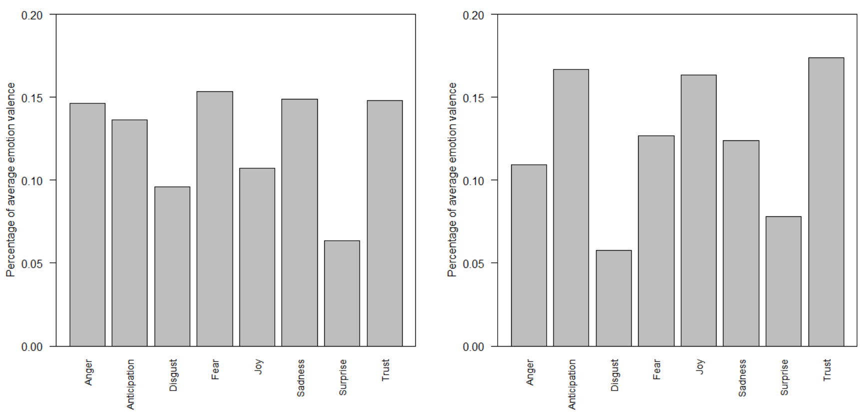

Figure 3

Percentage of the average valence of basic emotions in negative (left), and positive reviews (right)

In case of negative reviews, the most frequent emotions felt were fear (15.33%), sadness (14.87%), and trust (14.78%). Anger was a close fourth (14.64%). In case of positive reviews, the three most frequent emotions felt were trust (17.38%), anticipation (16.68%), and joy (16.34%).

When comparing all emotions between the groups, the Kolmogorov-Smirnov test was used (due to the large sample size) to assess their normality. Each was rejected ($p < 2.2\times10^{-16}$), therefore, they did not follow Gaussian distribution (since $p < 0.05$). Therefore, the nonparametric Wilcoxon rank sum test was chosen for the comparison. According to the results, all emotions were significantly different from each other between the two groups ($p < 2.2\times10^{-16}$).

Afterward, the sentiments found in the reviews were compared. They can be observed in Figure 4.

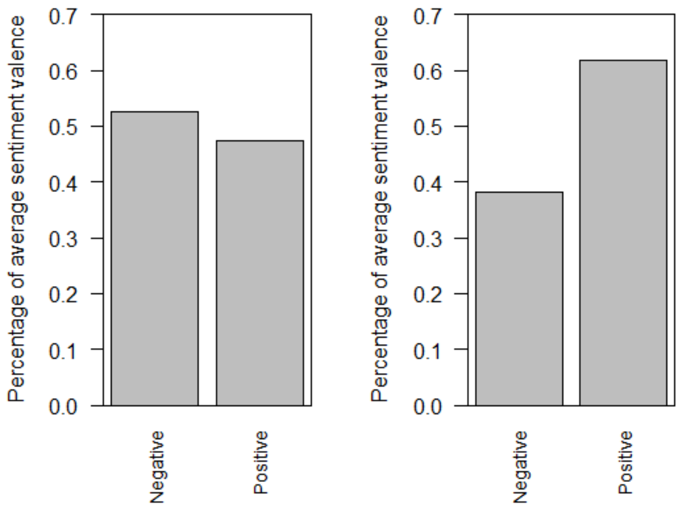

Figure 4

Percentage of average sentiments in negative (left), and positive reviews (right)

The percentage of negative sentiments' valence was smaller by 27.30% in positive reviews, and the percentage of positive sentiments' valence was larger by 30.27% in positive reviews. Their differences were also investigated with the Wilcoxon rank sum test (as the data distribution is still nonparametric). According to the results of the test, the differences between them were significant in each case ($p < 2.2\times10^{-16}$). In case of negative reviews, the difference between negative and positive sentiments was 0.0515494, while it was 0.2355724 when positive reviews are investigated. Although both differences are significant ($p < 2.2\times10^{-16}$ in each case). Clearly, the difference is much higher in the case of positive reviews.

3.2 Difference between Emotional Valence during Narrative Time

In this subsection emotional valence was investigated based on the passage of time. Imagine the following: when something is happening, for example, a player writes a review, time passes during the process. This phenomenon is called narrative time, or in other words, the time it takes to finish a text or a story. Using this phenomenon, it is possible to examine the emotional valence in each moment. In most cases, a point in narrative time could simply mean a sentence in the text. However, since the reviews have different lengths, each text was divided into an equal number of chunks, and the mean sentiment valence was calculated for each. 50 chunks were chosen on an empirical basis. Therefore, the narrative time of 1 means the first chunk of the text, while the narrative time of 50 means the last chunk of the text in this case. The results of this analysis are shown in Figure 5.

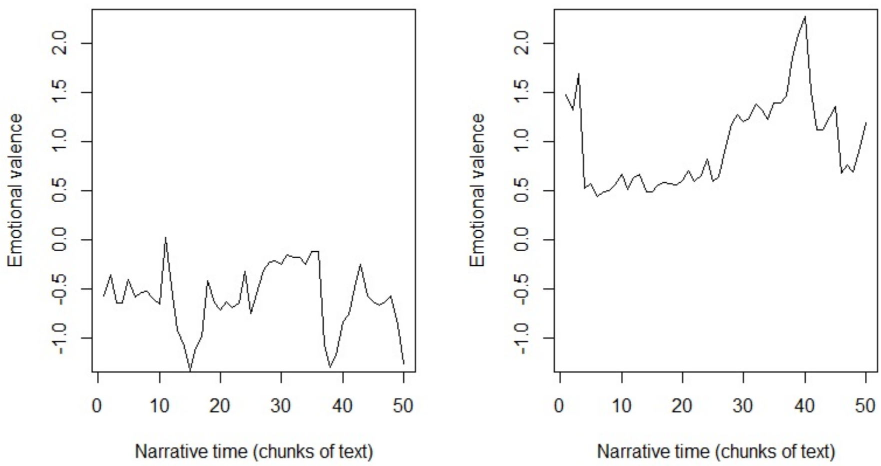

Figure 5
Emotional valence in negative (left), and positive reviews (right)

As can be seen, negative reviews mostly consisted of negative emotions, as the valence was below zero. There were times when some positive emotions could be found in the text, and that would increase the emotional valence (e.g. around the narrative time of 10). However, in case of negative reviews regarding challenging video games, the emotional valence did not exceed zero. Positive reviews were more interesting on the other hand: reviews usually started with a higher emotional valence, then it quickly decreased. Possibly, this was where the reviewers talked about the negative aspects of these video games. Afterward, the emotional valence quickly increased again. All in all, it can be easily observed that the emotional valences stayed around their starting values: the standard deviations were 0.32, and 0.45 in cases of negative, and positive reviews, respectively.

Lastly, they were compared to each other. Using the Shapiro-Wilk normality test, their distribution was assessed: $p = 0.1126$ in case of negative reviews, and $p = 0.0001686$ in case of positive reviews. Using the Wilcoxon rank sum test (due to the latter's nonparametric distribution), the following results were received: the two emotional valences were significantly different from each other as their means were -0.59, and 0.97, respectively ($p < 2.2\times10^{-16}$).

3.3 Difference in Commonly used Words

To understand the reasons behind the analyzed emotions, the word frequencies were counted in the reviews. Their frequencies were plotted into two word clouds, as can be seen in Figure 6. The larger the word, the more frequent it was. Besides this, 30 frequent words in negative reviews are observable in Table 4.

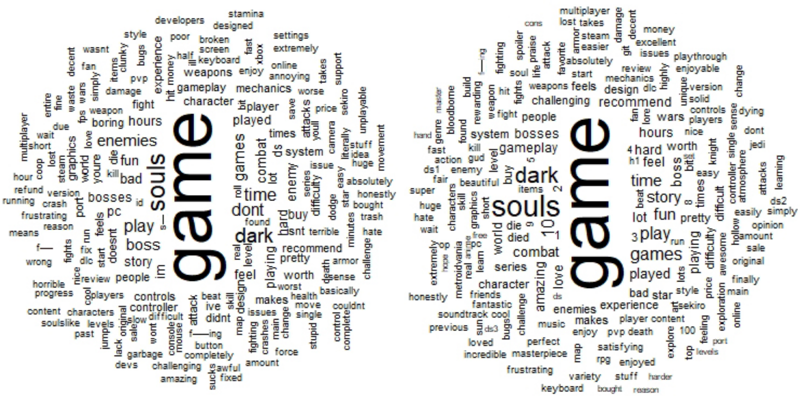

Figure 6
The most frequent words in negative (left), and positive reviews (right)

Table 3
30 frequent words in negative reviews

Game	92010	PC	6723	Frustrating	2113
Souls	24782	Hard	6002	Annoying	2043
Dark	18086	Feel	5895	Crash	1973
Play	14225	Hours	5799	Death	1952
Time	12760	Buy	5642	Worse	1875
Boss	10944	Level	5383	Waste	1756
Enemies	10942	Design	5220	Broken	1719
Combat	8738	Lot	5079	Poor	1661
Bad	8044	Hit	5074	Garbage	1545
Fun	7932	Die	4237	Unplayable	1426
Story	7335	Terrible	2753	Hate	1330

Naturally, in case of the negative reviews, more negative words could be found. Other negative words included the s-word with a frequency of 2823, and even the f-word could be found in two forms (2310 and 1616 times). The word “refund” appeared in 1575 sentences, while “refunded” occurred in 368. This suggests that some people were so angry with these types of games that they either considered a refund or already got one.

Due to the words, these two cases could have happened: the first is a hardware, software, or optimization problem. The words “crash”, “crashes”, “broken” and “poor” appeared in many sentences, which suggest that the game did not work properly. The second case is the difficulty of these types of games. As can be seen, the words “die”, “death” have high occurrences. The words “enemies”, “boss”, and “combat” were the 6th, 7th, and 8th most frequent words, respectively. Thus, the negative feelings of players arose from the challenge itself: their in-game characters died many times and they became frustrated or angry.

On the other hand, while positive reviews contain fewer negative words, they presented PX in more detail. Besides words of emotions, things that people liked about the games could be found in the reviews. 30 frequent words in positive reviews can be seen as examples in Table 4.

Table 4

30 frequent words in positive reviews

Game	450692	Love	31602	Worth	22260
Souls	146083	Gameplay	31119	Experience	22207
Dark	107760	Lot	30633	Challenging	20726
Play	61812	Hours	30221	Difficulty	19941
Fun	57943	World	29460	Star	19301
Story	54135	Feel	28313	Bit	19178
Time	53485	Recommend	27470	Wars	19132
Combat	45996	Enemies	26402	Level	19121
Boss	39752	Die	26340	Design	18373
Hard	35399	Pretty	23561	Character	18211
Amazing	33512	Buy	23276	Lore	15580

Other common words were “difficult” (17705), “system” (16708), “atmosphere” (8791), and “masterpiece” (8015). According to the words, players liked the high difficulty, combat with enemies and bosses, and the world of games. The latter can include level design and interaction with characters.

The number of words was also different between the two types of reviews. The average number of words in case of negative reviews was 113.12 with a standard deviation of 185.80. These numbers were 51.73 and 118.70, respectively in case of positive reviews. This means that overall, negative ones were longer and more detailed. According to the results of the Wilcoxon rank sum test (as this dataset was also nonparametric due to the results of the Kolmogorov-Smirnov test), the number of words was significantly different from each other as $p < 2.2\times10^{-16}$.

3.4 Difference in Agreeing on Reviews

In this subsection, the number of votes on reviews is analyzed. The number of votes on the reviews can be seen in Figure 7. It should be noted that reviews with zero votes are not plotted in Figure 7.

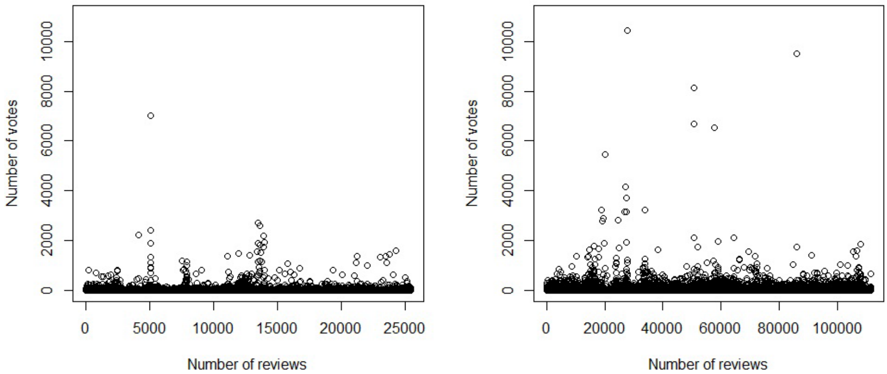

Figure 7
The number of votes on the reviews in case of negative (left), and positive ones (right)

In case of negative reviews, 25401 had at least one vote (out of 41149). This is 61.72% of them. Contrarily, 111537 had at least one vote (out of 377334) when positive reviews were investigated. This means 29.55% of the positive reviews. This can mean the following: people are more likely the agree with negative reviews as their experiences were similar. This is further strengthened by the fact that negative reviews have a mean of 6.81 and a standard deviation of 64.62 votes, while positive ones have 2.07 and 42.95 votes, respectively.

Lastly, the top five voted-on reviews were investigated in case of negative and positive ones. They can be observed in Table 5.

Table 5
The number of votes and words of the top five negative (left) and positive reviews (right)

	Number of votes	Number of words		Number of votes	Number of words
Negative 1	7038	238	Positive 1	10453	10
Negative 2	2700	13	Positive 2	9523	444
Negative 3	2593	19	Positive 3	8131	312
Negative 4	2417	589	Positive 4	6682	14
Negative 5	2232	9	Positive 5	6541	406

Across the data, many short reviews were voted up. From the textual analysis, it can be observed that these short reviews were usually funny or sarcastic sentences that presented games in a short form. However, the question may arise due to the variable number of votes and words: Does the number of words correlate with the number of votes? To answer this question, a correlation test was done in case of each group of reviews. As the data did not follow Gaussian distribution ($p < 2.2\times10^{-16}$ in case of both groups according to the Kolmogorov-Smirnov test), Spearman's rank correlation method was used and its results are presented in Figure 8.

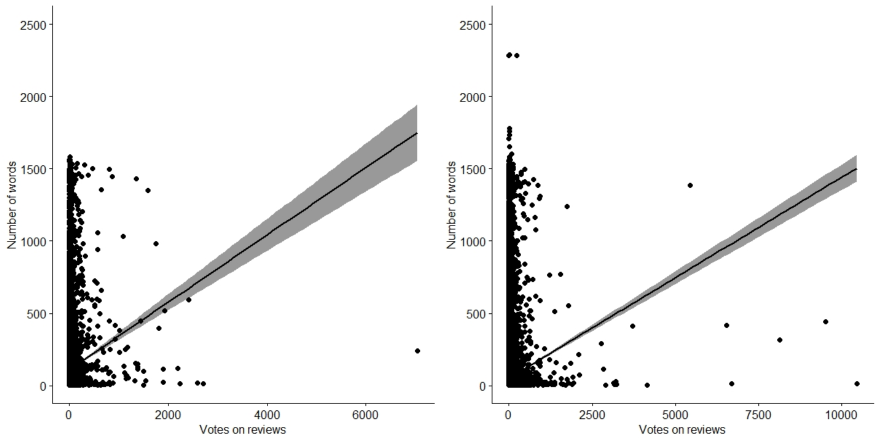

Figure 8

Correlation between the number of words and votes on reviews in case of negative (left), and positive ones (right)

According to the results received on the correlation test, no correlation can be found among the number of words and votes: $\rho \approx 0.2$ in case of negative reviews, and $\rho \approx 0.15$ in case of positive ones. Although both were significant with $p < 2.2\times10^{-16}$. Therefore, it can be concluded that the length of a review did not influence the number of votes. This means that players did not vote based on the length of the reviews.

4 Discussion

As PX is complex and can vary due to many factors, it is difficult to fully examine it. However, with the results presented in this article, it is possible to understand PX regarding challenging role-playing video games as well as the structure of both negative and positive reviews. The results may also help the designers of future games of the same genre.

According to the results, the structure of negative and positive reviews was significantly different. No transition exists between them: there were no negative reviews that included a large number of positive sentiments, and there were no positive reviews that had a large number of negative sentiments. It should be noted that in case of negative reviews, positive sentiments were much higher compared to the negative sentiments in positive reviews. The difference between negative and positive sentiments in negative reviews was 0.0515494, while it was 0.2355724 in case of positive reviews. Clearly, the gap between the two types of sentiments was much higher. To understand this gap, Figure 5 should be looked at to see the emotional valences of both review types. According to their, standard deviations, negative reviews did not deviate as much from their mean as positive ones. The standard deviation of the former was 0.32, while it was 0.45 of the latter.

Also, there were significant differences in case of all eight basic emotions between the two types of reviews. The emotions of anger, disgust, fear, and sadness significantly occurred fewer times in case of positive reviews. The decreases in frequency were 25.27%, 39.95%, 17.41%, and 16.61%, respectively. The feeling of disgust had the largest decrease in frequency, while sadness had the smallest one. Contrarily, the occurrences of emotions of anticipation, joy, surprise, and trust were significantly different between the two types of reviews. These differences were 22.12%, 52.26%, 22.85%, and 17.54%. In this case, joy was the most frequent, while trust was the least frequent. It can also be concluded that when a game was recommended, the frequency of joy was the largest, while when a game was not recommended, the frequency of disgust was the largest.

A similar phenomenon could be found when word frequencies were assessed. In case of negative reviews, players stated their disgust and anger towards games: they usually detailed software/hardware/optimization problems, and in some cases it was mentioned that games were difficult – impossible even. Due to the emotions of the player, the average length of negative reviews was 118.67% longer than positive ones. This means that positive reviews were shorter by 51.73 words on average. Therefore, positive reviews were less detailed. Even though they were shorter, the experiences regarding gameplay were more detailed. As the positive emotions were stronger, players focused on what they liked about games.

Even though the number of words was significantly different in case of both types of reviews, it can be concluded that they did not correlate with the number of people agreeing with reviews. This means that it was not the word number that mattered when agreeing with a review, but the content itself. In the end, people agreed with negative reviews more than with positive ones.

While multiple papers investigated Steam reviews from various perspectives, only a few could be compared to this paper. One such comprehensive study was written by Lin et al. [54]. They also did a thorough analysis of textual reviews. Although they compared reviews regarding several video game genres, not regarding multiple games in a subgenre. Still, it is possible to compare the results about review contents and length to their results. According to their research, negative reviews usually contain bug reports and complaints about game design, which is similar to the results presented in this study. Regarding review length, they concluded that negative reviews are only slightly longer than positive ones. The case is similar regarding “Souls-like” video games.

Conclusions

Understanding and investigating PX are quite complex, and the experience itself can vary due to different factors, such as the difficulty level of games. In this paper, the PX in case of challenging (or “Souls-like” to be more precise) video games was investigated by analyzing Steam reviews. The analysis was done in the statistical program package R and consisted of four investigations.

The first one investigated the difference in emotions and sentiments between negative and positive reviews. Significant differences were found among all eight basic emotions: anger, disgust, fear, sadness were significantly less frequent in case of positive reviews, while anticipation, joy, surprise, and trust significantly occurred more frequently. Regarding sentiments, significantly fewer negative ones, and significantly more positive ones can be found in positive reviews than in negative reviews.

In the second investigation, the emotional valence among both groups of reviews was examined. This was done with the help of the concept of narrative time. In case of negative reviews, the emotional valence was below zero, while it was above zero when positive reviews were investigated.

The third investigation focused on the word frequencies found in both groups of reviews. While negative reviews usually used words to detail software/hardware/optimization problems, positive ones often described player experiences or what they liked about games.

The focus of the fourth investigation was on the number of votes on and words in both negative and positive reviews, and whether a correlation could be found between them. According to the results, there was no correlation among them. However, negative reviews were longer and more likely to receive votes than positive ones.

In conclusion, by analyzing how negative and positive reviews are structured with experiences and emotions, it is possible to get a glimpse of how people felt when playing challenging games. Those who did not recommend a game (in other words: wrote a negative review) included some positive sentiments in the review, while even fewer negative experiences were included from those who recommended a game (in other words: wrote a positive review). Still, it is also possible that people are more likely to express positive opinions, but less likely to express negative ones, in which case they just agree with what others have written.

However, since about 90% of the analyzed reviews were positive, it can be concluded that the PX regarding these games is good overall.

Acknowledgement

This work has been implemented by the TKP2021-NVA-10 project with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the 2021 Thematic Excellence Programme funding scheme.

References

[1] T. Dixon: From passions to emotions: The creation of a secular psychological category. Cambridge University Press, 2003

[2] J. E. LeDoux, R. Brown: A higher-order theory of emotional consciousness. Proceedings of the National Academy of Sciences, Vol. 114, No. 10, 2017, pp. E2016-E2025

[3] O. Larue, R. West, P. S. Rosenbloom, C. L. Dancy, A. V. Samsonovich, D. Petters, I. Juvina: Emotion in the common model of cognition. Procedia computer science, Vol. 145, 2018, pp. 740-746

[4] I. Granic, A. Lobel, R. C. Engels: The benefits of playing video games. American psychologist, Vol. 69, 2014, pp. 66-78

[5] J. Scoresby, B. E. Shelton: Visual perspectives within educational computer games: effects on presence and flow within virtual immersive learning environments. Instructional Science, Vol. 39, No. 3, 2011, pp. 227-254

[6] J. Lessiter, J. Freeman, E. Keogh, J. Davidoff: A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. Presence: Teleoperators and Virtual Environments, Vol. 30, No. 4, 2001, pp. 282-297

[7] B. G. Witmer, M. J. Singer: Measuring Presence in Virtual Environments: A Presence Questionnaire. Presence: Teleoperators and Virtual Environments, Vol. 7, 1998, 225-240

[8] S. Björk, J. Holopainen, Patterns in game design. Cengage Learning, 2005

[9] D. A. Norman, S. W. Draper: User centered system design; new perspectives on human-computer interaction. L. Erlbaum Associates Inc, 1986

[10] W. Inchamnan, P. Wyeth: Motivation during videogame play: Analysing player experience in terms of cognitive action. In Proceedings of The 9th Australasian Conference on Interactive Entertainment: Matters of Life and Death, pp. 1-9, 2013

[11] P. Baranyi, A. Csapo: Definition and synergies of cognitive infocommunications, Acta Polytechnica Hungarica, Vol. 9, pp. 67-83, 2012

[12] G. Sallai, The cradle of cognitive infocommunications, Acta Polytechnica Hungarica, Vol. 9, No. 1, 2012, pp. 171-181

[13] P. Baranyi, Á. Csapó, Á. Balázs, P. Várlaki: An Overview of Research Trends in CogInfoCom, 18th International Conference on Intelligent Engineering Systems – INES 2014, IEEE Hungary Section, Tihany, pp. 181-186, 2014

[14] P. Baranyi, A. Csapo, Gy. Sallai: Cognitive Infocommunications, Springer International Publishing Switzerland (978-3-139-19608-4)

[15] A. Kovari: Human-machine communication from an industry 4.0 perspective and its relationship with education 4.0, New challenges and pedagogical innovations in VET & HE, 2018, pp. 637-647

[16] I. Heldal, C. Helgesen: The Digital HealthLab: Supporting Inderdisciplinary Projects in Engineering and in Health Education, Journal of Applied Technical and Educational Sciences, Vol. 8, No. 4, 2018, pp. 4-21

[17] A. Kovari: Study of Algorithmic Problem-Solving and Executive Function, Acta Polytechnica Hungarica, Vol. 17, No. 9, 2020, pp. 241-256

[18] E. Gogh, R. Racsko, A. Kovari: Experience of Self-Efficacy Learning among Vocational Secondary School Students, Acta Polytechnica Hungarica, Vol. 18, No. 1, 2021, pp. 101-119

[19] R. Demeter, A. Kovari: Importance of digital simulation in the competence development of engineers defining the society of the future, Civil Szemle, Vol. 17, No. 2, 2020, pp. 89-101

[20] C. Vogel, A. Esposito: Interaction Analysis and Cognitive Infocommunications, Infocommunications Journal, Vol. XII, No. 1, 2020, pp. 2-9

[21] J. M. Olaso, M. Inés Torres: User Experience Evaluation of a Conversational Bus Information System in Spanish, 8th IEEE International Conference on Cognitive Infocommunications, Debrecen, 2017

[22] B. Szabo, K. Hercegfi: Research questions on integrating user experience approaches into software development processes, 8th IEEE International Conference on Cognitive Infocommunications, Debrecen, 2017

[23] D. Geszten, B. P. Hámornik, K. Hercegfi: User experience in a collaborative 3D virtual environment: A framework for analyzing user interviews, 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), pp. 207-210, 2015

[24] A. Esposito, A. M. Esposito, M. Maldonato, C. Vogel: Differences between hearing and deaf subjects in assessing foreign emotional faces, 8th IEEE International Conference on Cognitive Infocommunications, Debrecen, 2017

[25] S. Ondas, L. Mackova, D. Hladek: Emotion Analysis in DiaCoSk Dialog Corpus, 7th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), Wroclaw, 2016

[26] A. Esposito, A. M. Esposito, G. Cordasco, M. Maldonato, C. Vogel, N. Bourbakis: Emotional faces of children and adults: What changes in their perception, 9th IEEE International Conference on Cognitive Infocommunications, Budapest, 2018

[27] T. Longobardi, R. Sperandeo, F. Albano, Y. Tedesco, E. Moretto, A. D. Di Sarno, S. Dell'Orco, N. M. Maldonato. Co-regulation of the voice between patient and therapist in psychotherapy: Machine learning for enhancing the synchronization of the experience of anger emotion: An experimental study proposal. 2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), pp. 113-116, 2018

[28] M. Horsfall, A. Oikonomou: A study of how different game play aspects can affect the popularity of role-playing video games. 2011 16th International Conference on Computer Games (CGAMES), pp. 63-69, 2011

[29] R. Daneels, S. Malliet, L. Geerts, N. Denayer, M. Walrave, H. Vandebosch: Assassins, gods, and androids: How narratives and game mechanics shape eudaimonic game experiences. Media and Communication, Vol. 9, No. 1, 2021, pp. 49-61

[30] P. Skalski, R. Whitbred: Image versus sound: A comparison of formal feature effects on presence and video game enjoyment. PsychNology Journal, 2010, Vol. 8, pp. 67-84

[31] P. Jagoda: On difficulty in video games: Mechanics, interpretation, affect, Critical Inquiry, Vol. 45, 2018, pp. 199-233

[32] A. Summerville, J. R. Mariño, S. Snodgrass, S. Ontañón, L. H. Lelis: Understanding mario: an evaluation of design metrics for platformers, In Proceedings of the 12th international conference on the foundations of digital games, pp. 1-10, 2017

[33] I. J. Livingston, L. E. Nacke, R. L. Mandryk: Influencing experience: the effects of reading game reviews on player experience. In International Conference on Entertainment Computing, Springer, Berlin, Heidelberg, pp. 89-100, 2011

[34] J. A. Bopp, E. D. Mekler, K. Opwis: Negative emotion, positive experience? Emotionally moving moments in digital games. In Proceedings of the 2016 CHI conference on human factors in computing systems, pp. 2996-3006, 2016

[35] S. Hedegaard, J. G. Simonsen: Extracting usability and user experience information from online user reviews. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2089-2098, 2013

[36] E. Guzman, W. Maalej: How do users like this feature? a fine grained sentiment analysis of app reviews. In 2014 IEEE 22nd international requirements engineering conference (RE), pp. 153-162, 2014

[37] I. Busurkina, V. Karpenko, E. Tulubenskaya, D. Bulygin: Game Experience Evaluation. A Study of Game Reviews on the Steam Platform. In International Conference on Digital Transformation and Global Society, Springer, Cham, pp. 117-127, 2020

[38] D. Lin, C. P. Bezemer, A. E Hassan: An empirical study of early access games on the Steam platform, Empirical Software Engineering, Vol. 23, No. 2, 2018, pp. 771-799

[39] Y. Yu, B. H. Nguyen, F. Yu, V. N. Huynh: Esports Game Updates and Player Perception: Data Analysis of PUBG Steam Reviews. In 2021 13th International Conference on Knowledge and Systems Engineering (KSE). pp. 1-6, 2021

[40] L. Eberhard, P. Kasper, P. Koncar, C. Gütl: Investigating helpfulness of video game reviews on the steam platform. In 2018 Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS), pp. 43-50, 2018

[41] H. N. Kang, H. R. Yong, H. S. Hwang: A Study of analyzing on online game reviews using a data mining approach: STEAM community data. International Journal of Innovation, Management and Technology, Vol. 8, No. 2, 2017, p. 90

[42] T. Guzsvinecz: The Correlation between Positive Reviews, Playtime, Design and Game Mechanics in Souls-like Role-playing Video Games, Multimedia Tools and Applications, Vol. 82, 2023, pp. 4641-4670

[43] X. Haotian: steam-reviews, PyPI, 2021. https://pypi.org/project/steam-reviews/. Accessed 10 April 2021

[44] R Core Team: A Language and Environment for Statistical Computing, 2018, http://cran.univparis1.fr/web/packages/dplR/vignettes/intro-dplR.pdf. Accessed 18 April 2021

[45] User Reviews – Get List (Steamworks Documentation), 2021, https://partner.steamgames.com/doc/store/getreviews. Accessed 14 September 2021

[46] M. Jockers: Package 'syuzhet', 2017, https://cran.r-project.org/web/packages/syuzhet. Accessed 12 September 2021

[47] K. Werder, S. Brinkkemper: Meme-toward a method for emotions extraction from github. In 2018 IEEE/ACM 3rd International Workshop on Emotion Awareness in Software Engineering (SEmotion), pp. 20-24, 2018

[48] S. Kiritchenko, X. Zhu, C. Cherry, S. Mohammad: NRC-Canada-2014: Detecting aspects and sentiment in customer reviews. In: Proceedings of

the 8th international workshop on semantic evaluation (SemEval 2014), pp. 437-442, 2014

[49] F. Bravo-Marquez, E. Frank, B. Pfahringer, S. M. Mohammad: AffectiveTweets: a Weka package for analyzing affect in tweets, Journal of Machine Learning Research, Vol. 20, 2019, pp. 1-6

[50] S. Mohammad, P. Turney: Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. In Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text, pp. 26-34, 2010

[51] T. Mailund: Manipulating data frames: dplyr. In R Data Science Quick Reference, Apress, Berkeley, CA, pp. 109-160, 2019

[52] J. Silge, D. Robinson: tidytext: Text mining and analysis using tidy data principles in R. Journal of Open Source Software, Vol. 1, No. 3, 2016, p. 37

[53] A. I. Kabir, R. Karim, S. Newaz, M. I. Hossain: The Power of Social Media Analytics: Text Analytics Based on Sentiment Analysis and Word Clouds on R. Informatica Economica, Vol. 22, No. 1, 2018, pp. 25-38

[54] D. Lin, C. P. Bezemer, Y. Zou, A. E. Hassan: An empirical study of game reviews on the Steam platform. Empirical Software Engineering, Vol. 24, 2019, pp. 170-207