Self-Serving Bias and Toxic Communication in Online Game Communities

By Christopher Mach, Orange Coast College
Mentor: Baxter Krug

Abstract

The rapid growth of online spaces has led to the development of new forms of communication between individuals in these communities. However, this has also resulted in the emergence of toxic behavior. Previous research has suggested that toxic communication may be highly personal in nature. To investigate this, we conducted a survey of various subreddit communities using a Likert scale to gauge respondents' self-serving bias and toxicity scores. The results showed a low correlation between self-serving bias and toxicity, suggesting that the personal element in toxic communication may be weaker than anticipated. These findings could be used to inform further research into the behaviors and differences in communication within online gaming communities.

Key Words: Video Games, Self-Serving Bias, Toxic Communication, Reddit Communities

Introduction

Imagine this scenario: an ordinary player sits down to play an online cooperative team game called League of Legends (LoL). As soon as the game begins, another player starts causing problems, trying to get the team to exit matchmaking because one of the team members has a history of poor performance. When that doesn't work, he randomly targets another player to harass and begins to throw racial slurs at them. Despite all of this, the game somehow manages to start, but it's clear from the beginning that it's not going to go well. The toxic player continues to sabotage the team and complain about his teammates in chat for the next 15 minutes. Eventually, a forfeit proposal is called and the team votes in favor, ending the game early. It's a truly toxic experience that leaves everyone involved feeling frustrated and upset.

Background

The rapid growth of online spaces has led to the development of new forms of communication between individuals in these digital communities. However, this has also led to the emergence of toxic behavior, where individuals engage in insults and belittlement of others. This is particularly prevalent in team-based online cooperative multiplayer game communities, where players compete against each other to achieve victory.

Research on toxic culture and deviant behavior in gaming communities has been a confusing area of study. In a recent article, Kowert (2020) attempted to catalog and outline the academic discussion on "dark participation" in games. Kowert defines dark participation as an umbrella term for all deviant behavior that takes place online, with behaviors that cause harm to others' health or well-being considered toxic. Additionally, the distinction between trolling and toxicity was established, with trolling referring to the intent to engage in dark participation and toxicity to the outcomes of such behavior.

Building from Kowert's (2020) work, this study will examine toxic behavior in online gaming communities, with a particular focus on the role of self-serving bias in driving such behavior. It is hypothesized that this bias may correlate to a tendency in individuals to display toxic behaviors. This study aims to explore this link in the context of online gaming communities. This research is important because toxic behavior can have negative impacts on the individuals who experience it, as well as on the overall health and functioning of online gaming communities.

Toxicity

“Toxic” and “toxicity” are broad terms that can generally be taken as any negative behavior (Kordyaka et al., 2020). In online gaming spaces, toxic communication is mainly studied in reference to methods of detecting toxicity in communication spaces where other players are able to communicate through text. Toxicity in communities can damage the reputation of that community. This particularly can be a problem for the growth of that group. If a group is perceived as toxic, this can inhibit the growth of the game and thus act as a barrier for revenue of game companies (Kordyaka et al., 2020).

In each of these communities, communication can be described as generally positive, negative, or neutral. A study by Ayushi Ghosh analyzed various gaming communities on online platforms found that different game communities would hold different proportions of sentiment in discussion. For example, the LoL community was found to have predominantly positive discourse in their community while another game called Minecraft was found to contain generally neutral discourse (Ghosh, 2021).

Discussions of detection were mainly analyzed in the context of text-based communication. Martens et al., tried to detect text-based toxicity by analyzing the frequency and sequence of key words (Martens et al., 2022). However, in online games, toxic behavior is not limited to text/verbal-based communication, but also nonverbal actions and gestures. Players can make “physical” actions in the game world that communicate toxicity, but interpretation of that toxicity would be contextual to that community. For example, in the game LoL repeatedly using an in-game command to have the player’s character “emote” and make gestures to the opposing team would communicate toxic behavior to other players without the use of speech or text. However, let it be noted that this is distinctly different from Kinesics. While the example given is a type of nonverbal toxic communication, toxicity can still be observed in other nonverbal forms that do not necessarily involve facial expressions or gestures. Another example is team self-sabotage by deliberately performing in a way that would put an individual's team at a disadvantage. Of course, detection of these sorts of behaviors is more difficult and highly contextual.

It is well known that intense or controversial topics can be a catalyst for toxic behavior (Almerkehi et al., 2019). In the context of video games, certain in-game events, such as defeating an enemy player or being killed by an enemy player, can trigger toxic behavior from other players (Martens et al., 2022). The anonymity of online spaces may also contribute to the negative disinhibition effect, which can further fuel toxic behavior (Kordyaka et al., 2020). This seems to be a common theme in literature, as Neto et al. (2017) also define toxic behavior as a negative response to a game event that causes anger and frustration, leading to harmful communication (quoted in Kordyaka 2020, p. 3).

However, it is important to note that not all players will react to these triggers in the same way, or at all. A study by Kordyaka et al. proposes that a combination of social cognitive theory, the theory of planned behavior, and the online disinhibition effect can provide a more holistic understanding of toxic behavior (Kordyaka et al., 2020). In short, players are toxic because it is a social norm relative to a game community, they already plan to engage in toxic behavior, and the ability to stay somewhat anonymous online makes individuals more at ease with participating in toxic behavior. Relating to social cognitive theory, the results of this study suggest that for some communities, actions that are considered toxic by one group may not be seen as toxic by another. Moreover, individuals do not necessarily exhibit toxic behaviors given the same social parameters as their peers in their communities. This suggests that toxicity is highly personal and may be influenced by other factors.

In order to prevent toxicity in online game communities, it is important to consider the individual factors that may influence toxic behavior. For the purposes of this study, toxic behavior will be defined as any negative behavior that harms the game community and its atmosphere, including harassment, flaming, trolling, and cheating (Kordyaka et al., 2020). Investigating the relationships between these factors and toxic behavior may

help us to develop methods for preventing toxicity in online games by seeking out a relationship to personality variables.

Self-serving bias

The self-serving bias is a psychological phenomenon that can be described as “judgements or interpretations of oneself, one’s behavior, and the behavior of others in ways that are favorable to the self, without requiring that such judgments be accurate according to some objective standard” (Blaine & Crocker, 1993). While this has been studied in relation to other aspects of human behavior in various contexts, there is a lack of research on the self-serving bias specifically in online game communities. Previous studies on the self-serving bias suggest that it can be measured on a scale and the degree of the bias varies from person to person. Additionally, community respondents to a micro study conducted prior to this study reflect varying degrees of self-serving bias across individuals in the same community (see Phase 1 in methods).

A study by Duval & Silvia looked at the self-serving bias in individuals who were assigned to groups with either a high or low probability of improvement on a task. They found that when self-awareness was high, those who expected improvement were more likely to attribute failure to themselves than those who did not expect improvement. Additionally, when the probability of improvement was high, highly self-aware individuals attributed more causality to themselves than those with low self-awareness (Duval & Silvia, 2002). That is to say, individuals had varying degrees of self-serving bias tendency when they held different levels of self-awareness.

The degree of the self-serving bias was affected by factors such as the perceived probability of improvement. If the self-serving bias is varied, it can be measured on a scale and compared to the tendency for toxic behavior in online game communities. This study aims to determine if there is a correlation between the self-serving bias and toxic communication in these communities.

This research project hypothesizes the following:

- H1: There will be a positive correlation between self-serving bias and toxic behavior.
- H2: The community alignment to qualitative measures will reflect the results of the average self-serving bias and toxicity scores.

Methods

This project takes a mixed-method approach to data acquisition and analysis. The combination of quantitative and qualitative data was gathered in three phases for a later analysis.

Phase 1

A micro study was performed on subcommunities on an online forum called Reddit. Reddit is an internet forum where specific topics are discussed in smaller subgroups or "subreddits." The study focused on two communities of team-based cooperative multiplayer game players: LoL and Valorant (Val). In discussions of these two communities, topics such as toxicity and individuals engaging in self-serving bias were common. The selection of these topics was based on the author's familiarity with and involvement in these communities. Despite the existence of self-serving bias prevalent in these community posts, a fair number of comments to these posts contain community members criticizing the original poster for diverting responsibility to external factors and suggest taking more personal responsibility to improve as a player. It is assumed that other communities similar to LoL and Val may also exhibit similar discourse.

Phase 2

A survey was designed to match the needs of game communities. The measures were constructed in a way that ensured the questions would be understood by most members of any game community that involved

cooperative team-based multiplayer modes. Most existing measures focused on toxicity in a text-based context or did not gather information relevant to this study. Since this study aims to assess forms of toxicity from nonverbal communication, unconventional measures were used. The survey consisted of five initial true-false statements for qualitative data, followed by 22 Likert scale questions on a scale of 1 to 7 (ranging from strongly disagree to strongly agree and from never to almost always) divided evenly to gauge the degree of self-serving bias and toxicity in subjects. To ensure that the questions in the unconventional measures were properly understood, the measures were tested in a small convenience sample of cooperative team-based multiplayer game community members to adjust the wording. After going through the IRB review process, the board determined that since measures did not collect any personally identifiable information and posed little to no risk to subjects, the study was exempted from IRB review.

Phase 3

A convenience sample was initiated with the final survey measures. The data were collected from the larger groups of cooperative team-based multiplayer game communities on Reddit. In particular, the game communities of LoL, Val, Overwatch (OW), Counter Strike Global Offensive (CSGO), and DOTA 2 were surveyed because they had large communities of players and served as industry examples of cooperative team-based multiplayer games. Some level of member participation was expected based on the size of each individual community. Reddit posts containing a link to the survey and a glossary of slang terms used in gaming communities were provided in each of the communities' subreddits.

Subjects who participated in the survey were asked to confirm on a consent form that they were US citizens and at least 18 years old. They were also informed that participation was completely voluntary. The data collected from the surveys were then copied onto Google Sheets and imported into IBM SPSS for linear regression analysis and inferential testing on the average results of each community's self-serving bias and toxicity scores. Each community was analyzed individually.

Results

Despite promising participation from the majority of surveyed communities, CSGO and DOTA 2 communities had 1 respondent each and will be omitted from this analysis. A summary of T/F from the three remaining communities of interest is displayed in Table 1. LoL community had n = 22 responses, Val community had n = 21 responses, and OW community had n = 16 responses. The first question (Q1) “Loser's queue exists,” question two (Q2) “Each game match is balanced,” question three (Q3) “I become tilted from bad matches,” question four (Q4) “Toxicity is a problem in this community,” and question five (Q5) “I flame other players.”

Question Number	League of Legends		Valorant		Overwatch	
	T	F	T	F	T	F
Q1	40.9%	59.1%	47.6%	52.4%	25%	75%
Q2	22.7%	77.3%	19%	81%	12.5%	87.5%
Q3	72.7%	27.3%	71.4%	28.6%	43.8%	56.2%
Q4	95.5%	4.5%	95.2%	4.8%	87.5%	12.5%
Q5	50%	50%	23.8%	76.2%	12.5%	87.5%

Table 1: True-false Responses for Online Gaming Communities

The results of the study indicate that the LoL data set has the weakest correlation coefficient between self-serving bias and toxicity (r = .234) among the three data sets. The Shapiro-Wilks test of normality showed that

p > .05, indicating that the data are normally distributed. The residual plot for the LoL data appears scattered, indicating that a linear model is appropriate. Based on responses from the community, the self-serving bias outliers were at 5.36. The average self-serving bias score (M = 4.00, SE = .131) and toxicity score (M = 3.46, SE = .264) were calculated considering the outliers. The results of Harman's One-Factor test showed evidence of Common Method bias (variance > 50%).

The Val data set had the strongest correlation between self-serving bias and toxicity (r = .653) among the three data sets. The Shapiro-Wilks test of normality showed that p>.05, indicating that the data are normally distributed. The residual plot for the Val data shows a linear model is appropriate. Based on responses from the community, the self-serving bias outliers were at 2.00 and 1.27. The average self-serving bias score (M = 3.79, SE = .219) and toxicity score (M = 3.08, SE = .276) were calculated considering the outliers. The results of Harman's One-Factor test showed evidence of Common Method bias (variance > 50%).

The results of the study indicate that the OW data set has a moderate correlation between self-serving bias and toxicity (r = .541). The Shapiro-Wilks test of normality showed that p > .05, indicating that the data are normally distributed. The residual plot for the OW data shows a linear model is appropriate. Based on responses from the community, the self-serving bias outliers were at 6.09. The average self-serving bias score (M = 3.91, SE = .208) and toxicity score (M = 2.88, SE = .313) were calculated considering the outliers. The results of Harman's One-Factor test showed evidence of Common Method bias (variance > 50%).

The sum of all data sets had a moderate correlation between self-serving bias and toxicity (r = .484). The Shapiro-Wilks test of normality showed that p > .05, indicating that the data are normally distributed. The residual plot for the combined data shows a linear model is appropriate. Based on responses from the communities, the self-serving bias outliers were at 1.27, 2.00, and 6.09. The average self-serving bias score (M = 3.91, SE = .107) and toxicity score (M = 3.16, SE = .158) were calculated considering the outliers. The results of Harman's One-Factor test showed evidence of Common Method bias (variance > 50%).

Discussion

The following sections will address research hypothesis H1 and H2. The section on correlation will serve to address H1 while the section on the True-False responses will serve to address H2.

Correlation

The results of the present study showed that there was a positive correlation among all communities. This supports the initial hypothesis that there would be a positive correlation. For each community, the Harman test showed common method bias, results are inflated. All communities' correlation between the self-serving bias and toxicity are not strong (r < .7). Functionally, there is little to no correlation between the two variables. At least from this result, it can be argued that responsibility diversion is not a trigger for toxicity. Due to the presence of common method bias, the correlational results for this study are inconclusive. Further study into the relationship between the self-serving bias and toxicity may be required.

Self-Serving Bias

LoL community holds the highest-self-serving bias score (M = 4.00, SE = .131) followed by the OW community self-serving bias score (M = 3.91, SE = .208) and the Val community self-serving bias score (M = 3.79, SE = .219). Interestingly, the Self-Serving bias scores gets higher as the age of the community gets higher. It could be communities become more self-serving the older they get. The game LoL released in 2009, OW in 2016, and Val in 2020. This could be a potential spot for further research.

Toxicity

LoL community holds the highest toxicity score (M = 3.46, SE = .264) followed by the Val community toxicity score (M = 3.08, SE = .276) and OW community toxicity score (M = 2.88, SE = .313). Interestingly, the two highest toxicity scores are from communities that are intellectual properties of Riot Games. It could be possible that titles Riot Games create attract a demographic that are more predisposed to toxic behavior. Exploring the demographic of players that are attracted to titles of a particular game companies may be a potential spot for further research.

Self-Serving Bias True-False

The responses to the first two questions provided some interesting insights. In response to Q1, a majority of participants in the LoL, Val, and OW communities (40.9%, 47.6%, and 25%, respectively) indicated that they believe such a feature existed. In response to Q2, a majority of participants in the LoL, Val, and OW communities (77.3%, 81%, and 87.5%, respectively) stated they felt some of their game matches were not balanced, as external factors affected the fairness of a match. Despite LoL having a higher self-serving bias score than Val, the Val community had a higher proportion of players agree to statements diverting responsibility of in game performance to external factors. This is especially surprising, given that the Val community held the lowest self-serving bias score from the communities tested. The results are particularly interesting, as it could be interpreted that these communities may be more willing to engage in various degrees of responsibility shifting. It may be that communities will rationalize their performance in different forms, with some explanations for their losses considered too extreme. For example, a community may feel that the existence of a "losers' queue," is an extreme justification for poor performance, but are more willing to justify that some matches are not balanced as a rationalization. Further exploration into the differences of each community's tendency to resonate with certain forms of self-serving bias may be a spot for further research.

Toxicity True-False

The last three True-False questions measure the tendency for toxicity in the three communities. In Q3, a majority of the participants in the LoL, Val, and OW communities (72.7%, 71.4%, and 43.8% respectively) indicated that they were negatively affected from bad match experiences with an exception from the OW community. In Q4, a majority of the participants in the LoL, Val, and OW communities (95.5%, 95.2%, and 87.5% respectively) agreed that they had toxicity problems in their communities. In Q5, the participants in the LoL, Val, and OW communities (50%, 23.8%, and 12.5% respectively) stated that they participate in “flaming” or insulting other players. The responses from Q5 showed that the LoL community displayed a greater willingness to acknowledge their involvement in the toxicity issues prevalent in their community. The True-False responses align with the average toxicity score for the tested communities. Specifically, they were more inclined to admit to engaging in toxic behavior towards other players via flaming, as compared to members of the Val or OW communities. A possible explanation for the observed difference in toxicity among these games could be attributed to the differing behaviors of players from different game genres. Val and OW are both first-person shooter games, while LoL is a third-person, multiplayer online battle arena game. Exploring the relationship between game genres and toxicity may be a potential spot for further research.

Post Study Observations

The survey supplied a spot for members of the community to leave additional comments they may wish to provide.

League of Legends (LoL)

Comments from the LoL community focused on the self-serving bias aspects of the survey. They were also the only community to do so as both the Val and OW community's comments focused on toxicity. One respondent

from the LoL community commented on Q1 that “Its not a question if it losers queue exists or not because its factual.” Despite this respondent's assertion of a losers' queue, general responses to Q1 suggest that the members of the LoL community are divided on the existence of this in-game feature. LoL community members were the only group to comment on such a question. This finding could be used to inform further research on the differences in self-serving behavior among different game communities. Future exploration into the topic may be a potential spot for further research.

Valorant (Val)

Respondents from the Val community focused on commenting about toxic behavior. Community members made an effort to distinguish and specify at what points they engaged in toxic behavior as well as categorizing the types of people that engaged in such behavior. For example, one respondent commented that they would only “flame back” towards another player if the other person was the first to initiate such behavior. Another player reported experiencing targeted toxic behavior due to their gender, and expressed the belief that people should not engage in such behavior under any circumstances and should instead “mute toxic teammates.” In a comment by another person, they mentioned that they felt more frustrated from “bad teammates” (those who are toxic, uninvolved, or do not use in-game voice chat) compared to “bad players” (those who perform poorly in game). The ways in which the Val community members perceive and respond to toxic behaviors can provide valuable insights into the social dynamics of online gaming communities and the challenges they face. It may also be useful to consider whether similar patterns hold true in other gaming communities. Exploring this topic could be a spot for further study.

Overwatch (OW)

The OW community contained responses on toxicity similar to responses from the Val community. One individual responded that they only acted toxic towards “those who are [already] toxic” and that even then they would not act “too toxic.” Another respondent also felt that “disabling chat should be easier for players who can't handle toxicity” implying that when toxicity is encountered, it would be better to mute problematic individuals instead of firing back. A particular individual said, possibly in jest, that they loved to “tbag other players” and that most of their kills would result in “tbags even if it means [they] died whilst tbagging.” Reflecting on how the Val community responded similarly, exploring communities within the same genre and toxic tendencies for those groups may be a spot for further study.

Limitations

The sample in this study was a convenience sample taken from a sub community of each game genre. The results can only be generalized to players who are both members of the game's respective Reddit communities and play the game. Common method bias is also a concern in the results since data was collected through self-report measures. Results may be influenced since all the data were collected in similar ways.

Conclusion

Suspicion on positive correlation between the self-serving bias and toxicity has shown to be true according to the data gathered. However, the values from the data gathered suffer common-method bias and thus the results are inflated. There is reason to suspect that there is little to no correlation between the self-serving bias and toxicity. Additional research may be required to ascertain what sort of relationship exists between the self-serving bias and toxicity. As previously mentioned, results of the study can only be generalized to members of each respective game's Reddit communities. Given the results of the study, it may be worthwhile to explore other avenues of individual traits and their relationship to toxic tendencies in game communities. Further study into other areas concerning individual game communities and their differences towards certain self-serving or toxic tendencies independently may bring interesting results.

Analysis of individual communities shows slight differences in the attitudes toward self-serving bias and toxicity. The True-False data, particularly results from Q5, indicate that the unique differences in each game community could explain their different reactions to the measures of the study. Differences in these behaviors could suggest that strategies to address toxic communication may need to be tailored for the specific characteristics of each community. Further exploration into the reasons for each communities' different responses and reactions to certain toxic behaviors could be a potential spot for further study.

References

[1] Almerekhi, H., Kwak, H., Jansen, B. J., & Salminen, J. (2019). Detecting toxicity triggers in online discussions. Proceedings of the 30th ACM Conference on Hypertext and Social Media. https://doi.org/10.1145/3342220.3344933

[2] Blaine, B., & Crocker, J. (1993). Self-esteem and self-serving biases in reactions to positive and negative events: An integrative review. *Self-Esteem*, 55–85. https://doi.org/10.1007/978-1-4684-8956-9_4

[3] Duval, T. S., & Silvia, P. J. (2002). Self-awareness, probability of improvement, and the self-serving bias. *Journal of Personality and Social Psychology, 82*(1), 49–61. https://doi.org/10.1037/0022-3514.82.1.49

[4] Ghosh, A. (2021). Analyzing Toxicity in Online Gaming Communities. *Turkish Journal of Computer and Mathematics Education, 12*(10). https://doi.org/10.17762/turcomat.v12i10.5182

[5] Kordyaka, B., Jahn, K., & Niehaves, B. (2020). Towards a unified theory of toxic behavior in video games. *Internet Research, 30*(4), 1081–1102. https://doi.org/10.1108/intr-08-2019-0343

[6] Kowert, R. (2020). Dark participation in games. *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.598947

[7] Martens, M., Shen, S., Iosup, A., & Kuipers, F. (2015). Toxicity detection in multiplayer online games. *International Workshop on Network and Systems Support for Games (NetGames)*. https://doi.org/10.1109/netgames.2015.7382991

Appendix A

Glossary

- Flaming – Sending angry, derogatory, or inflammatory messages/comments to other players.
- Tbag – the act of crouching and standing in repeated cycles on a dead player's corpse. It is often used as a method to mock, belittle, or humiliate the opponent.
- Losers' queue – An alleged in-game mechanic with a bias that places the individual in a team of other players that are statistically likely to lose so that a victory for an individual is incredibly difficult to achieve.

Survey Measures – Self-Serving Bias

From a likert scale 1-7...

1. I lose games because of things out of my control
   Strongly Disagree – Strongly Agree
2. I suffer from bad luck in games ____
   Never – Almost Always
3. My equipment holds me back from winning games
   Strongly Disagree – Strongly Agree
4. I alone control if I win games
   Strongly Disagree – Strongly Agree
5. Luck is a factor in people’s ranks
   Strongly Disagree – Strongly Agree
6. My teammates make winning games impossible
   Strongly Dsiagree – Strongly Agree
7. I find myself in losers' queue
   Never – Almost Always
8. Some matches can't be won
   Strongly Disagree – Strongly Agree
9. Bad teammates lose me games
   Never – Almost Always
10. Unbalanced characters lose me game
    Never – Almost Always
11. I play against teams with better players
    Never – Almost Always

Survey Measures – Toxicity

From a likert scale 1 – 7...

1. My rank should be higher
   Strongly Disagree – Strongly Agree
2. Most people deserve a lower rank
   Strongly Disagree – Strongly Agree
3. It is okay for me to flame my teammates if they troll
   Strongly Disagree – Strongly Agree
4. Throwing an unwinnable game is okay
   Strongly Disagree – Strongly Agree

5. Some people deserve to get trolled
   Strongly Disagree – Strongly Agree
6. If someone is being flamed, they probably deserve it
   Strongly Disagree – Strongly Agree
7. People should not play online games if they can't handle being flamed
   Strongly Disagree – Strongly Agree
8. It is okay to flame bad players
   Strongly Disagree – Strongly Agree
9. It is okay to be toxic if the other person started it
   Strongly Disagree – Strongly Agree
10. I'm better than most players in my rank
    Strongly Disagree – Strongly Agree
11. I find myself becoming tilted from bad players
    Never – Almost Always