Toxicity by Game Design: How Players Perceive the Influence of Game Design on Toxicity

ZINAN ZHANG, The Pennsylvania State University, USA
SAM MORADZADEH, The Pennsylvania State University, USA
ANDREW WOAN, The Pennsylvania State University, USA
YUBO KOU, The Pennsylvania State University, USA

Toxicity in online games refers to behaviors where players disrupt the gaming experience of others, leading to adverse outcomes such as depression and low self-esteem. Although scholars have identified various factors contributing to toxicity, ranging from individual motivations to team dynamics to cultural backgrounds, the role of game design has been less frequently discussed. To bridge this gap, we conducted an interview study to explore players' perceptions of how game design influences toxicity. Our research identified four game design elements that participants perceived as contributing factors to the emergency of toxicity in their experiences: team interdependency, fairness, interaction design, and privacy. These findings help us shed light on how game design unintendedly triggers toxic intentions, exposes players to vulnerability, making them potential victims, and affects player interactions which lead to toxicity. We further propose design implications that can mitigate toxicity in online games.

CCS Concepts: • **Human-centered computing** → Human computer interaction (HCI) → Empirical studies in HCI

KEYWORDS: Toxicity, Game Design, Video Games, Online Games, Design Affordance, Unintended Consequences

ACM Reference format:

Zinan Zhang, Sam Moradzadeh, Andrew Woan and Yubo Kou. 2024. Toxicity by Game Design: How Players Perceive the Influence of Game Design on Toxicity. Proc. ACM Hum.-Comput. Interact., 8, CHI PLAY, Article 345 (October 2024), 31 pages. https://doi.org/10.1145/3677110

1 INTRODUCTION

Toxic behavior is an umbrella term to describe "behaviors that disrupt players' morale and team dynamics" [73]. It has various forms such as griefing [24], trolling [16], sexual harassment [11], general harassment, flaming, cheating, and cyberbullying [47]. Toxic players have various

Authors' addresses: Zinan Zhang, zzinan@psu.edu, The Pennsylvania State University, E342 Westgate Building, University Park, Pennsylvania, USA; Sam Moradzadeh, szm6500@psu.edu, The Pennsylvania State University, E342 Westgate Building, University Park, Pennsylvania, USA; Andrew Woan, ajw6289@psu.edu, The Pennsylvania State University, E342 Westgate Building, University Park, Pennsylvania, USA; Yubo Kou, yubokou@psu.edu, The Pennsylvania State University, E353 Westgate Building, University Park, Pennsylvania, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
2573-0142/2024/10 - 345
© Copyright is held by the owner/author(s). Publication rights licensed to ACM.
https://doi.org/10.1145/3677110

reasons for engaging in toxic behavior. Sometimes, toxicity is intentionally conducted by these players [66]; sometimes it is a consequence of inadequate support in coping with stressful conditions in gaming [34]; and sometimes the toxic players are not familiar with community norms and thus unintentionally violate the norms [45]. Toxic behavior is widespread, with recent surveys and studies revealing a significant percentage of participants (some over 80%) experienced some form of harassment in online multiplayer games [46, 84–86]. Even worse, yearly consecutive studies by the Anti-Defamation League find this number increasing with every year [85]. Exposure to toxicity causes many players to suffer from severe, negative, and sometimes lasting consequences like lower self-esteem, increased anxiety, depression, and decreased enjoyment with the game and outside of the game [34, 63, 82, 85]. Consequently, many players either stop playing games altogether or resort to other methods, such as hiding their identity [84].

To understand toxicity and address its adverse effects on players, game researchers have committed to explaining what may cause toxicity. For instance, researchers have highlighted individual motivations [24, 29, 39, 51] as a major factor. For instance, Liu and Agur [51] suggest that players might conduct toxic behavior due to intrinsic motivations, such as prioritizing assisting their real-life friends in gameplay while disregarding others, which can be seen as toxic by other players. Researchers have also traced toxicity back to its social context [3, 5, 16, 42]. For example, Kou [42] highlights that social dynamics, like conflicts within teams, can foster toxicity. Some have also probed how culture affects one's behavior and thus fosters toxic conduct [29, 32, 34, 35, 41]. For example, Kordyaka et al. [35] argue that cultural differences among players can lead to varying levels of toxicity. As game scholars explore what constitutes toxicity, they also acknowledge how toxic behavior is subjective and influenced by games' capabilities, ethics, rules, and norms [34, 42]. Kou [42] describes this as a "situated" view of toxicity. Kowert [46] takes a similar view, discussing how trash-talking may be considered fun rather than toxic during an eSports tournament.

Scholars have touched upon how certain game design factors might inadvertently encourage toxicity. For example, Kou [42] identifies specific design aspects, like communication features that facilitate verbal abuse and the competitive nature of games like League of Legends, as potential enablers of toxicity. Türkay and Adinolf [72] point out that the option to add friends in Hearthstone could lead to verbal abuse. Deslauriers et al. [20] observe that game design elements such as repetitive tasks and role identification may contribute to toxic behavior in Dead by Daylight. Similarly, Burnay et al. [11] report that the portrayal of sexualized female characters in Ultra Street Fighter IV can incite sexual harassment. Kordyaka et al. [37] recently have noted that game design and its affordances significantly influence player interactions in general online games. However, it remains underexplored as to what are overarching design patterns that act as catalysts for toxicity across different online games.

This paper seeks to bridge this gap by examining how players perceive the influence of game design on toxicity across different online games. While players are not game design professionals, their extensive experiences with toxicity can provide valuable perspectives on the relationship between game design and player toxicity. Thus, we pose the following research question:

How do players perceive the influence of game design on toxicity?

To address this question, we conducted interviews with 20 individuals who have encountered or conducted toxicity in online games. Through an inductive thematic analysis [15] of the interview data, we pinpoint four game design patterns that can incite toxic behavior,

including 1) team interdependency, referring to design patterns that affect the reliance among teammates and heighten the stresses of collaboration; 2) fairness, related to design patterns that impact players' perception of equity and can create situations perceived as unfair to certain players; 3) interaction design, concerning design patterns that facilitate toxic interactions between players; 4) privacy, about design patterns that risk exposing players' personal information in ways they may not consent to. These findings allow us to discuss how game design influences toxicity from the standpoint of the toxic perpetrators (i.e., the player who conducts toxic behavior), the victims, and the affordances that contribute to toxic conduct. While the specific features we discuss may not apply universally across all online games, understanding the overarching patterns from these features and their impact is crucial.

Our study underscores a design perspective to understand toxicity: how game design influences toxicity. This broad view helps uncover underlying design patterns linked to the emergence of toxicity. Our contributions include 1) providing overarching game design patterns players perceived that can influence toxicity across different online games, 2) proposing a conceptual framework of the relationship between game design and toxicity, and 3) offering actionable insights for game designers to mitigate toxicity through designs.

In the subsequent sections of this paper, we first examine previous studies focusing on toxicity attribution in online gaming, as well as the impact of game design on toxicity. Following this, we elaborate on the methodology for our study, including data collection and data analysis. Subsequently, we elucidate four game design patterns that influence toxicity. Drawing from these insights, we propose a framework and offer design implications to mitigate toxicity. Lastly, we outline the limitations and future directions of this study before concluding with a summary.

2 RELATED WORK

Toxicity has various forms [16, 24, 25, 47] and can happen due to multiple reasons [34, 45, 66]. This study is not limited to a specific form or reason of toxicity. Instead, this study aims to understand how game design influences general toxicity in online games, specifically from the players' perspectives, examining both how it contributes to and mitigates toxicity. In this section, we first explore previous scholarly discussions on the attribution of toxicity in online games, including various factors such as game design. Then, we delve into how scholars have previously discussed the ways in which game design can also positively influence toxicity from players' perspectives, specifically how players perceive game design's role in mitigating toxicity.

2.1 The Attribution of Toxicity in Game Research

To effectively tackle toxicity and its adverse effects on players, it's crucial to understand what causes toxicity. This understanding can equip researchers and practitioners with the knowledge needed to identify the root causes of toxicity, thereby enabling them to develop targeted strategies for its reduction. So far, game research has identified several key factors contributing to toxicity, including individual, social, cultural, and game design factors [37, 43, 51, 57]. Individual factors, such as personality [29, 37] and motivation [24, 39, 51], play part of the role in toxic behavior. For instance, players with aggressive personalities or less pro-social tendencies are more likely to engage in aggressive and potentially toxic actions [18]. Similarly, those motivated by the desire to disrupt gameplay for others (known as griefing) also contribute to the toxic environment [24].

Social factors also influence toxic behavior. Specifically, team dynamics in games where players are unfamiliar with each other can often become a hotbed for toxicity [43]. The absence of strong bonds among team members makes it difficult to cultivate a positive and supportive atmosphere within the team. This environment frequently leads to diminished accountability and a tendency to blame one another for losses, facilitating toxic interactions [3, 5, 16, 42].

Cultural factors, including the cultural backgrounds of players and the prevailing norms within the gaming community, also play a role in fostering toxicity. Differences in behavior and attitudes towards toxicity can often be traced back to the cultural context of the players. For instance, players from cultures that emphasize collective social roles and expectations, such as India, may exhibit different patterns of behavior in games compared to players from North American cultures, where individualism is more pronounced [41]. This divergence may influence how players perceive and take responsibility for their actions within online gaming environments. Furthermore, the normalization of toxic behavior within the gaming culture itself perpetuates a cycle of toxicity, where toxicity spreads as victims retaliate against the perpetrator [38]. As toxic behavior becomes more accepted or expected within these communities, it encourages the further spread of such behavior, thereby exacerbating the issue [29, 32, 34].

We refer to game design as not only design features such as communication channels, but also broader level mechanisms such as game events' timing and criteria for winning [2]. Scholars discussed how game design factors can influence toxicity. At an abstract level, game design which emphasizes competition can override other valuable aspects like cooperation, friendship, and community, contributing to a toxic gaming environment [62]. Competitiveness in game design can also easily trigger conflicts within a team and thus facilitates toxicity [42]. Anonymity in game design influences one's perception about the consequences of conducting toxicity (e.g., lack of consequences for toxic behavior) and thus can contribute to toxic conduct [6, 14, 34, 51].

In addition to how game design causes toxicity, scholars found certain game design features have correlations with toxicity as well. For instance, games with sexualized characters are correlated with instances of sexual harassment [11], players controlling characters with higher kill counts tend to exhibit more toxic behavior [50], and violence within games is positively associated with toxicity occurrence as well [48]. Research has also delved into the affordances of game design to understand how they facilitate toxic behavior. For example, players use the 'add as friend' feature to write comments to other players from previous matches to harass them [72]; players spam others in games using the emote feature to humiliate and annoy them [72]; and players engage in verbal abuse through team communication channels [64].

In sum, the complexity of attributing toxicity arises from the interplay of individual, social, cultural, and game design factors. While considerable research has explored the first three factors [3, 5, 16, 24, 29, 29, 32, 39, 41, 42, 51], attention to game design factors has only recently increased. In addition, most of the scholars have investigated how certain game design features correlate with toxicity or how these features cause toxicity in a specific game. There is a notable gap in research efforts dedicated to explicating a broader view of how game design patterns influence toxicity across a wide range of online games. This paper aims to address this gap by exploring how players perceive the relation of game design patterns to toxicity.

2.2 Game Design for Mitigating Toxicity

Design has the potential to guide human behavior in a specific direction, which, if not carefully considered, can veer towards immorality [31], and thus may lead to toxicity. However, game design can also play a productive role in mitigating toxicity. In order to make a moral gaming environment, before implementing a design, designers must propose how players might react to it [1]. To aid in this process, Flanagan et al. [22, 23] proposed a framework of game design, values at play (VAP), urging game designers to create environments where morality is designed in the game, and players are encouraged to engage in moral behavior.

Several studies have proposed game designs to guide players towards moral behavior (i.e. non-toxic behavior) in gaming environments [19, 36, 49, 54, 77]. For example, Munn [54] proposed to encourage prosocial activities and positive behaviors through badges, rewards, or other items. Kordyaka and Kruse [36] suggested implementing a respective game design to dissociate perceived toxicity and educate players to maintain a friendly gaming environment. Depping et al. [19] proposed to have asymmetric abilities in character and team setup design to reinforce the importance of collaboration and encourage friendly communication. All these designs aim to promote a friendly gaming environment and moral behaviors, thereby mitigating toxicity.

As toxicity can negatively affect players' experiences and well-being, some research has explored designs aimed at reducing its adverse impacts on players' wellness. For example, Wu et al. [77] suggested that game design can support victims by using their experiences as a teachable moment for them to deal with loss and frustration. Reid et al. [63] suggested to integrate in-game support tools for immediate and post-game recovery (e.g. eye bleach pictures and seeking for warm messages after exposure to toxicity). While these designs do not decrease the incidence of toxicity, they offer support to victims affected by toxic behavior, thereby alleviating the adverse effects of toxicity within gaming communities.

Addressing toxicity through game design is crucial for creating a welcoming gaming environment, as well as promoting moral behaviors and supporting players' well-being, especially from players' perspective. As a human-computer interaction (HCI) researcher, it is crucial to incorporate users' viewpoints into analysis and design. This notion has further propelled HCI design methodologies, such as participatory design [67, 68], which prioritize incorporating users' viewpoints into the design process. While game design is conventionally seen as a domain primarily involving game designers, there appears to be a tendency to downplay the roles of players from the game design perspective. Although numerous design strategies have been proposed by researchers to tackle toxicity, only a few studies have explored the viability and acceptance of these measures from the players' perspective (e.g., [63]). This oversight could lead to game designs that inadequately consider players' perspectives. Despite players not being professional game designers, a comprehensive examination of their views of how game design may influence toxicity, in either positive or negative ways, can yield valuable insights on game design to mitigate toxicity accordingly.

3 METHODOLOGY

In this paper, we conducted a qualitative study to understand **how players perceive the influence of online game design on toxicity**. The perspectives from players can offer a nuanced players' view of the relation between game design and toxicity, allowing for reflective evaluation of past designs and informed guidance for future design choices. To answer the

research question, we conducted 20 semi-structured interviews, followed by a thematic analysis. The study was approved by the university's IRB office prior to the data collection.

3.1 Data Collection

We recruited our participants via the first author's personal contact (n=2), snowball sampling (n=3), and recruitment messages posted on our university campus (n=15), a large, public university based in the U.S. Snowball sampling is a widely employed method in qualitative research [55, 60]. It relies on network and referral-based participant recruitment, starting with a small set of initial contacts and expanding through recommendations from existing participants, ultimately aiming to achieve a target sample size or saturation point [60]. Participants were requested to complete a pre-screening survey to provide demographic information prior to scheduling an interview session. Our participants were aged between 18 and 30. All of them have played at least one online game and have experienced toxicity or conducted toxicity (i.e. what they perceived as toxic). Except for two participants recruited from the first author's personal contact, all other participants were from the same institutions as the authors. All participants were students. For other demographic information such as race, gender, and age, please refer to Table 1 below. Within the pre-screening survey, participants were also asked to list all the online games they have played. While during the interview, we did not specifically ask about the games they play, we observed that some participants, when describing their experiences of toxicity, referenced games not previously mentioned in the pre-screening survey. Consequently, we added these additional games to the table below as well.

The first author conducted and recorded all the interviews via Zoom with the participants' consent. With semi-structured interviews, we were able to ask open-ended and follow-up questions that encouraged participants to discuss their toxicity-related experiences of online games in a natural flow without a structured sequence [17]. This method allowed new ideas to emerge from conversations with participants [17]. We conducted all interviews between February 2023 and July 2023, each lasting between 30 minutes and an hour. Due to variations in players' willingness to express and the number of details they can remember from past experiences, the interview time varies. Our interview guideline mainly included the following interview questions: 1) participants' demographic information and their general experiences with online games (e.g. what online games they have played); and 2) participants' experiences with toxicity, including their experiences either as a victim or as a toxic perpetrator, their explanation of toxicity, and their perceptions regarding the role of game design when confronted with a specific toxic behavior.

All the interview audio was transcribed using Zoom. We stored and analyzed interview transcripts and coding data in Word and Excel on password-protected computers. We compensated each of the participants with a $20 Amazon Gift Card after their interviews. To protect participants' privacy, we refer to them using code names (i.e., P1, P2, P3, ...).

Table 1. Participants' demographic information. The recruitment column represents how the participants were recruited. UC represents University Campus; SS represents Snowball Sampling; PC represents Personal Contact. Game column shows all online games participants have played. We used abbreviations to keep the table concise. LoL represents League of Legends, CoD represents Call of Duty, HoK represents Honor of Kings, DBD represents Dead by Daylights, OW represents Overwatch, CS:GO represents Counter-Strike: Global Offensive, Apex represents Apex Legends, WoW represents World of Warcraft, FF represents Final Fantasy, PUBG represents Player Unknown's Battlegrounds, GTA 5 represents Grand Theft Auto V, CR represents Clash Royale, TFT represents Teamfight Tactics in League of Legends.

No.	Age	Race	Gender	Recruitment	Game
P1	20	Asian	Male	UC	LoL, Apex, Master Duel, PUBG, OW
P2	20	Asian	Male	UC	Valorant, CoD, CS:GO
P3	18	Asian	Male	UC	LoL, Valorant, HoK, Rainbow 6, CS:GO
P4	23	White/Caucasian	Male	UC	LoL, OW, FIFA, CS:GO
P5	24	Black/African American	Male	UC	CoD, Fortnite, Apex
P6	21	White/Caucasian	Male	UC	Valorant, CoD, CS:GO
P7	21	Mixed	Male	UC	LoL, OW
P8	20	White/Caucasian	Female	UC	CoD, Minecraft, Roblox
P9	21	White/Caucasian	Nonbinary	UC	Fortnite, GTA 5, CoD
P10	22	Asian	Male	SS	OW, DBD
P11	23	Asian	Male	PC	OW, Rainbow 6
P12	23	Asian	Female	PC	OW, HoK
P13	25	Asian	Male	UC	LoL, CR, Hearthstone, PUBG, Fortnite
P14	29	Mixed	Male	UC	Halo series, CoD, Gears
P15	30	White/Caucasian	Female	SS	Fallout 76, WoW, Guild Wars 2, FF 14
P16	23	Asian	Female	UC	LoL, Minecraft
P17	26	Asian	Female	UC	LoL, PUBG
P18	26	Asian	Female	SS	TFT, Werewolf (Langren Sha)
P19	25	Asian	Female	UC	HoK
P20	23	Asian	Male	UC	LoL, CS:GO, Apex, Battlefield, Minecraft, Dota 2, Mario Kart

3.2 Data Analysis

We conducted an inductive reflexive thematic analysis (RTA) [15] of the interview data based on our RQ, aiming to understand how players perceive the game design's influence on toxicity. Two researchers analyzed data together. According to six-phases of RTA [9], they first reviewed each transcript to understand preliminary ideas on the relation between game design and toxicity and held a first-round discussion with three authors. This phase helps researchers familiarize themselves with the data. Then, both of them independently went through every transcript and coded quotes that were related to game design and toxicity. This step generated

over 453 initial codes. After the coding phase, the next phase is to generate initial themes. In this phase, three researchers conducted meetings to discuss their codes and resolved disagreements iteratively to reach a final set of initial codes. They finally agreed on a list of 69 second-level codes for the RQ. In the next phase, two researchers discussed several rounds to develop higher-level themes in the codes to capture the mechanism behind the game designs until they reached an agreement on a thematic map, which met the principles of internal homogeneity and external heterogeneity in thematic analysis [61]. The fifth phase asks researchers to refine, define, and name the theme. In this phase, two researchers discussed several rounds to settle done the theme names to finalize the thematic map. In the final phase, two researchers selected the most representative quotes carefully and documented them with theme names and codes in a finding outline with Word.

Pertaining to principles in RTA, we did not calculate inter-rater reliability [7]. This is because, in RTA, themes do not underlie the data, waiting to be discovered with objective agreement. Instead, the quality of coding stems from an in-depth engagement with the data and is situated in reflexive interpretation. Because of this, it is also typical to have a single coder for RTA [8]. However, our RTA was performed by two researchers. Additionally, the concept of RTA results in that data saturation is hard to align. In RTA, the key is to develop themes theoretically, where new meanings are always possible. Therefore, this development process can continue indefinitely. Consequently, in RTA, it is the researchers' responsibility to determine whether the goal is met with the current data and thus decide when to stop the analysis. A pragmatic practice for RTA is to determine a sample size. Our sample size is consistent with commonly accepted standards in the HCI community [12].

The resulting thematic map included four overarching themes, each representing a design pattern: team interdependency, fairness, interaction design, and privacy. We are inspired by these terms' definitions from previous work in different contexts [33, 58, 59] and define each design pattern in the Findings section to align this work's unique context. As an example of how initial codes were categorized into one of the overarching themes, the quote 'Smaller team might be more toxic than others because the longer you play… Each person's responsibilities responsibility is bigger than larger teams, so their behavior will directly affect how the team performs if the team was small' was initially coded under the theme 'smaller team setting.' After discussion, we had an agreement that even though, from a surface level, the participant perceived that toxicity is influenced by the game design of a smaller team, the reason behind it is the tight collaboration between teammates in the smaller team setting. Therefore, we coded it under the second-level sub theme 'intra-team tensions by intense teamwork design' under the first high-level theme 'team interdependency.'

4 FINDINGS

In this section, we describe how participants perceived the influence of game design on toxicity. There are four game design patterns that participants believe can affect toxicity: team interdependency, fairness, interaction design, and privacy. For each game design pattern, we detailed the concrete game designs participants mentioned in a diverse range of games.

4.1 Team Interdependency

Team interdependency refers to the extent to which players must rely on their teammates in order to advance or succeed in games. This interdependency intensifies with certain game designs, potentially leading to conflicts and, consequently, toxicity among players. Our

participants specified two types of game design patterns that can contribute to the emergence of toxicity.

4.1.1 Intra-Team Tensions by Intense Teamwork Design

Some participants pointed to the demanding nature of teamwork in competitive games as a root cause of toxicity. The need for precise coordination among teammates fosters a tense and pressurized environment within the team. In such circumstances, the subpar performance of even one team member can detrimentally affect the team's overall performance and the result of the game. This can lead to frustration among teammates who have put significant time and effort into the game, thus acting as a catalyst for toxic behavior, such as verbal abuse. A participant pointed out the substantial impact an individual's performance can have on team dynamics in games where close teamwork is essential, which in turn can escalate toxicity.

> *"Smaller team is more toxic than others... Each person's responsibility is bigger than larger teams, so their behavior will directly affect how the team performs if the team was small... In Rainbow 6, there are five players in a team... So, each player has a lot of responsibilities, like searching on the map to find enemies and kill them. If one player did not search well and died, the team would lose a teammate in the rest of the time [which would create a huge disadvantage for the team]." P11*

Rainbow Six is a tactical first-person shooter game in which players engage in realistic, close-quarters combat, with an emphasis on strategy, teamwork, and the precise execution of planned assaults against threats. P11 noted that in a small team, such as a 5-player team in Rainbow Six, each player carries a significant portion of the responsibilities to contribute to the team's success, more so than in larger teams. Consequently, any slight deviation from an individual's responsibilities can lead to a substantial disadvantage for the team. P11 observed that players are often more critical of someone who causes a significant disadvantage in a small team setup, leading to a tension of in-team gaming environment and thus tendency towards toxicity. Such toxic behavior tends to intensify, especially under the pressures of a competitive environment. For example, P14 shared their experiences in Gears 5:

> *"[In Gears, toxicity] normally occurs in the Horde mode because it takes hours. If you all die once on the mode, you have to start from scratch. Players normally weed out by level 10, but if you get up into the 30s and 40s, people aren't getting out. They will yell at you for not doing something right or blame you for being out of position. The longer you're in it, the more you want those achievements." P14*

Gears 5 is a third-person shooter video game that emphasizes cover-based combat, cooperative gameplay, and a narrative-driven campaign. In Gears 5, Horde mode is a game mode where a team of players battles through 50 waves (i.e., levels) of enemies' (non-player characters) attacks. To win, players must survive all 50 waves. If all players die, they have to start over from the first wave, rather than the wave where they were defeated. Such game design for win intensifies the tension between teammates. As P14 noted, the competitiveness in this game design intensifies as the levels go up and consequently increases the toxicity. This is because as the level increases, players are more afraid of losing and starting from scratch. This mindset amplifies the competitiveness and pressure to collaborate. Therefore, players under pressure are easily triggered and proceed to conduct toxic behavior (e.g., verbal abuse) to the player who has made a single mistake.

Conversely, participants also described how toxicity levels differ in games that do not have intense teamwork. This difference in toxicity is attributed to a design where an individual's

behavior has less influence on the performance of other players or the team as a whole. In such settings, where one's actions have minimal impact, toxicity tends to be less likely. For instance, a participant mentioned:

> *"Battlefield is nice because there are four people in a group. Whenever someone leaves the group, there will be a new player added to your group right after someone." P12*

In many video games, leaving before the game ends is considered toxic and is reportable, as these games typically require players to stay and compete until the end. Leaving a game early in such video games results in a team being short-handed, creating a disadvantage for the team and negatively affecting the experiences of its players, thereby rendering it toxic behavior. However, in Battlefield's setup, such behavior is not considered toxic. Battlefield is a first-person shooter video game series known for its large-scale multiplayer battles, combining infantry combat, vehicle warfare, and a class system that encourages teamwork and strategy across expansive maps. In these large-scale multiplayer engagements, it's common for multiple teams to compete (i.e., there are two sides, and each side has multiple teams to fight for the same goal). Players have the flexibility to switch teams, leave the game early, or join during an ongoing match. Within this game design, P12 highlighted how the act of a player 'leaving the game' is perceived as inconsequential to both individual and team performance due to the availability of new players to join the team. This ongoing replacement of players diminishes the significance of leaving during the game and reduces the associated tension in collaborative game design. Thus, the game's design alters player perceptions of leaving early from a toxic action to a non-issue.

4.1.2 Conflicts by Imbalanced Role/Character Design

In many team-based video games, players are assigned to different roles and characters, each with unique responsibilities and capabilities for teamwork. This design of asymmetrical role [28] in the team makes the collaboration between teammates important and thus intensifies team interdependency. While teammates are supposed to fulfill responsibilities and capabilities they are expected to do, the imbalanced role and character designs can lead to incapability of fulfilling the expectations and thus leads to conflicts in gameplay, which can trigger verbal abuse and other toxic behaviors. A participant explained how the Jungle role in League of Legends (LoL) can lead to such conflicts, with the potential to trigger toxicity:

> *"[In LoL,] Jungler on the other side has ganked you a lot of times... This is pretty boring and when you cannot farm and could provide a bad feeling. Then you are going to call your Jungler like why you didn't gank my lane. And if the Jungler says I'm ganking on the other line... this will create a war." P3*

League of Legends (LoL) is a competitive multiplayer online battle arena (MOBA) game where two teams of five players each strive to destroy the opposing team's base while defending their own on a map of three lanes and jungle areas in between. The five players typically fill out five distinct roles. The Jungle role (Jungler) involves moving through these jungle areas to support the rest of the team in three lanes, also known as 'ganking.' Players often expect the Jungler to visit their lane and assist them by ganking. However, as P3, a player who has 8 years of expertise in playing MOBA games, described, Jungler may not be able to gank every lane as other players' expectation, because there are three lanes and a Jungler can only gank one lane at a time. The player who does not receive the help of a gank may get triggered and blame the Jungler for not meeting their expectations, resulting in toxicity (i.e. a 'war' as P3 mentioned, referring to verbal abuse to each other). The toxicity, while seemingly a

result of Jungler's personal in-game decision (i.e., choose a lane to gank), partly stems from the design of the Jungle role, which entails significantly more collaboration with all other teammates. In other words, the design of workload for different roles is imbalanced, leading to much more burden on the Jungler role. Therefore, in a game with such role designs, toxicity is more likely to happen around the players in this role. Moreover, not just the role but also character design can result in different expectations of one's responsibility in teamwork and thus leads to toxicity:

> *"In Overwatch, there are some healers who are more likely to be blamed by their teammates, like Moira and Zenyatta... players in these characters don't heal but try to make damage..." P10*

Overwatch is a team-based multiplayer first-person shooter game where players are assigned a distinct role and select from a diverse cast of characters, each with unique abilities, to complete objectives and battle against the opposing team on various maps. The role in Overwatch represents the character type who has certain capabilities, such as Support, who has abilities to heal teammates, or Tank, who can take damage from enemies. In Overwatch, a team typically includes two Support roles, who are supposed to aid teammates (e.g., heal) rather than dealing damage. However, as P10, a player who spends 10 hours per week in playing Overwatch, highlighted, certain Support characters in Overwatch are uniquely crafted with dual capabilities allowing players to attack and heal. These dual capabilities lead to ambiguity in character roles, fostering divergent interpretations and consequently diverse gameplay actions. Such variations may not align with team expectations and thus can foster toxicity. For instance, as P10 noted, while teammates anticipate Support roles to provide assistance, players in these characters might instead engage in combat (i.e., do damage). This discrepancy between expectations and actual gameplay actions can spark conflicts and act as a trigger for toxicity. This toxicity arises not merely from players' actions, but also from the inherent ambiguity in character design that offers players varied gameplay options. These ambiguous designs draw attention to the players' gameplay choices, particularly when specific character capabilities are only tied to certain roles (e.g., only characters in Support role can heal). This situation can exacerbate toxicity around players in these characters who did not heal as expected. In summary, toxicity in gameplay often emerges from a clash of expectations, influenced both by the assigned responsibilities of each player's role and the versatile, yet unclear, capabilities of characters within the game.

The unclear responsibility and capability may not only align with the character's ability but also with the team composition. For example, a participant explained how toxicity was triggered by failure to assign responsibility when selecting the character in LoL:

> *"...[In LoL,] it's hard to play especially during the team fight, because if the enemy team has like 2 tank and three damage, while we only have like 5 damage but no tank, then it's like even if we do everything well like eventually, we'll all die and lose the game... they actually ruin each other like, saying why didn't you pick the tank champ when you saw we don't have any tank champ. Like they blame each other about the champion selection... In champ select stage, if one person selects something and if it can show like the AP and AD balance or like the tank and like damage balance, or the win rates of it. Then I think people can actually be a lot more cautious about what champ they select." P16*

In LoL, players are assigned a role and can pick a character from more than 160 characters to fill the role. But the characters have different types. For example, a character who can deal significant damage is an AP or AD character type, and a character who can withstand

considerable enemy attacks is a tank character type. P16, who played LoL for 18 months, posits that a balanced team composition, featuring a mix of all character types, could foster more enjoyable and winnable gameplay. Nonetheless, players might not always select character types that contribute to a balanced team, potentially diminishing the team's likelihood of victory despite collective efforts. This issue of character type selection stems from ambiguities in team responsibilities and the distribution of responsibilities of picking the character types, as dictated by the game's design. The ambiguities stem from players' varied interpretations of what constitutes a balanced team composition, and the absence of a clearly defined team composition of responsibility in character type selection. As a result, five players could eventually make an imbalanced team composition, creating disadvantages that may result in defeat, and, sometimes, toxicity within teams as players verbally abuse each other for not adhering to expected responsibilities regarding character selection balance. Thus, P16 even expressed a desire for a design feature that could provide information about whether their team composition is balanced.

4.2 Fairness

Some participants identify fairness in design as a contributing factor to toxicity. Fairness in game design means the extent to which a game design can provide every player with an equal chance (i.e. symmetrical fairness [59]) to compete and enjoy the game. Fairness perception matters to player experience. A game design that results in an unequal chance of enjoying the game can foster toxicity. According to participants, some of the game designs foster toxicity because they do not seem to provide an equal chance for every player to enjoy the game. For instance, a participant shared:

> *"CS:GO is a 5V5 shooter game and you can team kill in that game. So sometimes only four of my friends and I are online. We would get some random fits and try to communicate with him or ask him to do something. If he didn't do it, we would shoot him, and he would die... You can only kill your teammates twice per person. So, after we'd kill him the max amount of times, we do a vote to kick the player from the game." P6*

Counter-Strike: Global Offensive (CS:GO) was a first-person shooter game that sets two teams against each other, Terrorists and Counter-Terrorists, in various objective-based modes, focusing on team strategy, precise shooting, and map knowledge, with gameplay emphasizing bomb planting/defusal and hostage rescue missions. In CS:GO, players have the option to vote to kick a player if they perceive them as a hindrance to gameplay (e.g., toxic players). This option to vote is designed to empower players to make a democratic decision to remove toxic players and thus maintain a playful environment. However, when a group of players bands together, they have disproportionate influence during the voting process. This collective influence creates an unfair situation and could foster toxic conduct if their influence is exploited intentionally. For instance, as P6 stated, if a single player does not align with the desires of a four-player group (in a five-player team), that group might vote to kick the dissenting player out, regardless of whether that player's behavior was actually toxic or deserving of such action. In such scenarios, the vote of the individual player becomes insignificant, as the group of four can manipulate their numerical advantage, creating an imbalance in group dynamics. This toxicity lies in the design that empowers the majority group in the voting system, giving the minority group an unequal chance in the voting process.

Game design can also foster toxicity through unfair situations in the process of picking characters in the game. A participant shared:

> *"In LoL, it used to be you don't select a role... you can say I want mid but that doesn't guarantee that you get mid because the players ahead of you have priority on picking the roles they want... sometimes this causes toxicity because some people [ahead of you] would still pick the role they want... So you could see there' be fighting in the in the chat. It's not an official rule, just like a convention." P4*

In previous versions of LoL, the game requires players to select distinct roles during the role selection phase, with five roles available for the five players. This phase requires players to choose their roles sequentially, in a predetermined order. This arrangement creates an unfair situation where the player at the top of the order can choose their preferred role, whereas those later in the sequence have fewer opportunities to do so. Even though players could use the chat function to declare preferred roles, players with earlier selection opportunities may choose to ignore those declarations. Such behavior, as P6 claimed, incited arguments and, subsequently, toxicity. While toxicity appears to stem from the decisions made by players at the top of the order, it's also attributable, in part, to the game design, which grants them more power in role selection.

The last game design that fosters toxicity through unfair situations is the design of the encounter of players. In some online games, the encounter is dependent and regulated by the matchmaking system, but in some open-world video games, the encounter is unregulated and results in unfair situations:

> *"Previously in WoW, someone who is a higher level comes in and kills all the resources in the area because they're farming for themselves. And to me, that's kind of toxic cause you obviously see these low-level people need these things to level up and you're just farming for materials for your own stuff..." P15*

World of Warcraft (WoW) is a massively multiplayer online role-playing game (MMORPG) where players embark on quests, explore vast fantasy worlds, and interact with other players in a shared online universe. In previous versions of WoW, there were different level areas where all players shared the same resources and the same space. These spaces allow the encounter of high-level players and low-level players. P15 noted that high-level players may engage in toxic behavior by entering lower-level areas to monopolize resources, taking advantage of their higher-level status. This behavior creates an unfair situation for lower-level players since it leaves them resourceless: low-level players cannot access resources from other areas due to their inability to compete in higher-level areas. As a result, low-level players lose the opportunity to fully enjoy the game, with all accessible resources being consumed by higher-level players. The toxicity stemming from high-level players is a result of the game's encounter design, which creates an unfair situation for low-level players. This occurs because players of significantly differing levels are allowed to compete for the same resources within the same area, impacting the enjoyment of the game for those at lower levels.

4.3 Interaction Design

Interaction design is a design that suggests how players can use it to act in a multiplayer game and thus influence their fellow players' in-game experience. The previous two sections explained how game design patterns can affect interactions between players, which leads to toxicity. In this section, participants specified how interaction designs afford toxicity, allowing toxic interaction to happen. In other words, although the features are not designed for conducting toxicity, players reinterpret and appropriate them for toxicity. We found two types of affordances, affordance for communication and affordance for behavioral interactions.

4.3.1 *Affordance for Verbal Abuse in Communication Design*

Games provide players with multiple ways of communicating with each other via communication design. These designs may be exploited by players for toxic conduct, such as verbal abuse. A participant provided an example of verbal abuse in Gears 5:

> *"Gears was getting complaints about the language and verbal abuse. So, they stopped it for the newest game [Gear 5] and they kept it [text chat] off for the whole run of it... But players can still send messages to other players after the game... And it does feel more intentional, because they take the time to write you a letter now... It just seems much more passive as you weren't yelling at me during the game, so you felt the need to write me a note... and you can't have the actual reactions." P14*

In previous versions of Gears, there was an in-game communication channel. However, as P14 explained, game designers removed the channel to reduce verbal abuse in-game. Consequently, in the latest version, Gears 5, players no longer have an in-game communication channel. Instead, they are provided with an option to send messages after the game to players with whom they have previously played with. While this new communication design might reduce toxicity during the game, as P14 implied, they still experienced toxicity through messages received after the game. The impact of this post-game toxicity can be even more severe, as it often comes across as passive-aggressive and leaves P14 without the ability to respond directly, potentially leading to more profound psychological impacts on them. Therefore, the affordance in communication design, whether in-game or post-game, can serve as catalysts for enabling or facilitating toxicity such as verbal abuse.

4.3.2 *Affordance for Behavioral Abuse in Communication Design*

Certain interaction designs, whether advertently or inadvertently, serve as catalysts for abusive behavior during gameplay. A participant described how they engaged in toxic behavior through 'team damage,' which allows a player to inflict damage on their teammates, even to the point of killing them:

> *[Q: have you ever initiated toxic behavior in the game before and what's the reason?] "If some players intentionally sabotage gameplay, not try to win, or feed the team or something else, I will kill them just for making the team game easier." P11*

Team damage is a feature found in many first-person shooter games that enhances realism by simulating real-life shooting scenarios. However, this design also permits players to intentionally inflict damage on and even kill teammates, which can be considered toxic as P11 describes. This is because players who are killed have their gaming experience disrupted by the consequences of their death: they are present in the game but due to death, are unable to participate actively in gameplay. The ability to inflict team damage, although not designed for intentionally killing teammates, allows players to conduct such actions that can hinder the experiences of others, thus influencing toxic behavior.

Not only does disrupting others' chance to enjoy gameplay but also irritating others through sensory aspects of gameplay can disrupt gaming experiences and be considered toxic. A participant shared how players use soundtrack game design to annoy others:

> *"[In LoL,]...there is like some fancy music [when they dance]. So, I've seen a lot of times where people just like keep pressing it nonstop. So, you hear the noise all the time. It was very loud and annoying." P18*

In LoL, characters can dance accompanied by unique background music, intended to add humor and make light-hearted interactions with other players, including the enemies. However,

as P18 mentioned, some background music is loud, and if played repeatedly, can become annoying and thus be considered toxic. The act of dancing, while not directly disrupting others' gameplay, can interrupt the overall experience of other players, including the sounds they hear in the game. The sound for dance, although designed to provide fun player interactions, can be used in a toxic way.

Not all affordance in design leads to toxicity, some of them can help mitigate it. A participant shared how certain design affords a lower likelihood of players conducting toxic actions toward others:

> *"It's funny cause Gears does have in game gestures and you're allowed to like set up your character with like to actually give the middle finger, but no one does it well... You get 4 gestures in your gesture wheel. And people want the celebration gesture... or limited-edition ones... so they don't have a space for that [toxic gesture on their gesture wheel]. I think the main reason is that your character will actually pause and do gestures. There are multiple middle finger emotions, but that makes your character to stop and the other players actually have to be looking at your character. So, unless you want the middle finger to be your victory pose, it's not really a good method to communicate in the game." P14*

In Gears, players can do gestures, some of which, like the middle finger gesture, could be seen as disrespectful and potentially used in toxic manners. However, according to P14, players often opt not to use such gestures. This decision is influenced by two main factors. Firstly, players have limited space on their gesture wheel, a feature that allows them to select and customize which gestures they wish to have available during a game. Given this limitation, players prefer to reserve space for more valuable gestures, such as celebratory or limited-edition ones, rather than for gestures intended to provoke others. Secondly, the design of executing and seeing a gesture in Gears further discourages the use of disrespectful gestures. For a gesture to be seen, a player must actively stop and perform the gesture while other players need to be looking at them. As P14 suggests, the likelihood of this occurring during gameplay is low, making such gestures less impactful and, consequently, less appealing to players to have on their gesture wheel. Although toxic gesture designs exist, the limited capacity of the gesture wheel and the conditions required for gestures to be seen discourage players from utilizing them. Therefore, players tend to not conduct toxic behavior through these gestures, demonstrating how affordance in gesture design can influence player interactions and promote a less toxic gaming environment.

4.4 Privacy

Some participants mentioned how a lack of privacy protection in game design can contribute to toxicity. Privacy is defined as one's ability to control personal information about oneself for what information can be seen by others in which ways [58]. In this section, we found that the lack of privacy protection leads to a situation where information some players deem as personal is exposed to others unwantedly and is used in an inappropriate way after exposure to others, resulting in toxicity. For example, some female players discussed how their gender information could be exposed to other players via voice channel, compared to text channel, facilitating sexual harassment:

> *"I find it more enjoyable if I just don't turn on my microphone. So, if I'm playing, I'll use chat instead. I have a pretty gender-neutral username, so no one really knows that I'm a girl with just my username." P8*

> *"If they know you're a female and they just start talking to you like flirting, which is very annoying." P18*

Both P8 and P18, who are female players, have discovered that revealing their gender through the voice channel leads to unwanted flirting and even sexual harassment from other players, diminishing their enjoyment of the games and fostering a toxic environment. P8 has found that their gaming experience improves when they opt not to use the voice channel, which prevents others from discerning their gender based on their voice. For both P8 and P18, their privacy expectation is not to leak their gender information to other players, but voice channel violates such expectation. While what P8 and P18 shared could be attributed to sexism in the game culture, the design of voice channel also inadvertently serves as a means for disclosing gender identity through voice, compared to text-based channel (e.g., P8 would use 'chat' instead).

Besides gender identity, there is other information that, although players themselves may not want to share with others, is available to all players anyway due to the game design. For example, a participant shared how the design of performance statistics in game could be used to conduct toxic behavior:

> *"[In Overwatch,] it shows a lot of information including damage, accuracy, healing amount or tank amount... So, like if you play, let's say if you play a healer role and then you're not doing well, anyone can press tab and see your healing numbers like if it is too low and then they will [start abusing you]." P1*

In Overwatch, the game is designed to show everyone's performance statistics (e.g., how much damage one has made) in real time during a game. These statistics can unfortunately serve as ostensibly reasonable grounds for criticizing someone for their poor performance. As P1 mentioned, teammates may target the support player for their low healing numbers throughout the game. While these statistics might reflect performance to an extent, the statistics may not necessarily reflect all the performance aspects. However, players are frequently targeted simply because of their lower performance metrics. While gameplay-related statistics and visualization are highly valuable, to what extent such information should be shared with other players is worth thinking about. The lack of attention to privacy protection in this scenario, can orient players' attention to numbers, and serve as a catalyst for toxic behavior targeted at players with lower numbers.

Besides information on in-time performance statistics, another participant shared how a lack of protection on other's character preferences can facilitate toxic behavior:

> *"[In LoL,] I hate playing support. I like mid lane. If I don't get to mid lane, I'll ban others' characters... then they usually dodge like that's the whole point." P7*

In LoL, upon entering a match, players proceed to the character selection stage, where they can see their teammates' character preferences. Subsequently, the system prompts every player to ban one character they dislike playing against before selecting their own characters. P7 admitted that they conducted toxic behavior, banning teammates' preferred characters to trigger them intentionally. This behavior is based on the information of one's preference of character selection. While the visibility of the information on teammates' preference data may serve the function of preventing the case of unintentionally banning teammates' preferred characters, lack of protection of such information could let toxic players intentionally ban others' preferred characters based on the visibility of such information, leading to toxicity.

Finally, a lack of protection of information on a player's match history could also facilitate toxic behavior. A participant detailed how this information could be used in a toxic manner:

> *"[In LoL,] when you try to get promoted to an upper rank, you need to play three games and you need to win two of them. If [the toxic player] sees someone in the game who is on that promote. He [the toxic player] usually tries to throw the game so that that person can't go up to the next Rank." P16*

In an earlier version of LoL, players were required to succeed in a promotion series by winning two out of three games before they could advance to the next rank. As P16 noted, players' rank information, including details on their promotion status, was accessible online. This information allows for predictions on how the results of upcoming matches could influence one's rank. As such, toxic players, upon accessing this information, might intentionally lose games to sabotage another player's chances of winning their match and achieving promotion to the next rank. A lack of protection on information on promote status and previous match result information can trigger toxic perpetrators to conduct toxic behavior toward the targeted victim.

5 DISCUSSION

We reported how players perceive the link between toxicity and game design and identified four key design patterns (Table 2) within game design that could potentially influence player experiences and inadvertently foster an environment conducive to toxicity.

It is important to note that we are not suggesting that those games mentioned in our study are created with the intention of fostering player toxicity. Rather, we propose that toxicity is an unintended outcome of game design practices, which is distinct from games intentionally crafted to include manipulative features, often referred to as dark patterns [80]. The unintended outcome can be reflected in the behavior of players who have intentions of engaging in toxic actions, players who experience stress in gaming but lack adequate coping strategies, or players who are unfamiliar with the community norms. For example, players may intentionally sexually harass those from minority genders, as seen with P8 and P18. P11 can kill a toxic teammate as a coping strategy to confront toxicity. Some players may be unfamiliar with what others expect, in other words, the community norms in the game, which results in 'toxic' behavior, as P10 mentioned, trying to cause damage instead of providing support.

Game can be seen as a technology and unintended consequences of technology are not uncommon to see. Unintended consequences of technology have been recognized since the 1990s [71], and in various sectors like healthcare technology [4, 27], digital communication technologies [56, 65], and educational technology [70]. Unintended consequences may arise from corporate practices prioritizing shareholder interests [30] or being used by users in manners not foreseen by its designers [13]. As such, it is possible that even well-intentioned game design can become a contributing factor to toxicity, due to player interactions with and interpretations of the design. Methodologically speaking, prior research has employed methodologies like interviews to explore how user interactions with technology can result in unforeseen outcomes [13, 52, 74]. In a similar vein, our study employed semi-structured interviews to probe players' perceptions of what game designs could have toxicity as an unintended consequence.

Table 2. Summary of theme names, design patterns, and how design patterns trigger toxicity.

Theme Name	Design Pattern	How It Triggers Toxicity
Team Interdependency	The extent to which players must rely on their teammates in order to advance or succeed in games. For example, in a team-based game, players rely on collaboration in certain roles to win the game.	The interdependency intensifies with certain game designs, potentially leading to conflicts and, consequently, toxicity among players. For example, the design of the workload for different roles is imbalanced, leading to much more collaboration burden on one role. Therefore, in a game with such role designs, toxicity is more likely to happen around the players in this role.
Fairness	The extent to which a game design can provide every player with an equal chance (i.e. symmetrical fairness [59]) to compete and enjoy the game. For example, the design of encounters between high-level players and low-level players can result in losing accessible resources of low-level players, which creates an unfair opportunity for low-level players to play the game.	A game design that results in an unequal chance of enjoying the game can foster toxicity. For example, high-level players take the resources of low-level players, because players of significantly differing levels are allowed to compete for the same resources within the same area. Such behaviors impact the enjoyment of the game for those at lower levels and thus can be seen as toxic.
Interaction Design	A design that suggests how players can use it to act in a multiplayer game and thus influence their fellow players' in-game experience. For example, a player can dance and play music in games. Fellow players can see the dance, hear the music, and choose to respond to the player. The dance design is the interaction design.	Although the interaction designs are not designed for conducting toxicity, players reinterpret and appropriate them for toxicity. For example, the sound for dance, although designed to provide fun player interactions, can be used in a toxic way --- repetitively playing during the game.
Privacy	One's ability to control personal information about oneself for what information can be seen by others in which ways [58]. For example, a player may not want their gender information to be exposed to fellow players in the game.	Lack of privacy protection can lead to a situation where information some players deem as personal is exposed to others unwantedly and is used in an inappropriate way after exposure to others, resulting in toxicity. For example, the voice channel inadvertently serves as a means for disclosing gender identity through voice and triggers toxicity (e.g. sexual harassment) towards a specific gender.

5.1 Conceptualizing the Relationship between Toxic Behavior and Game Design

We acknowledge that toxicity can originate from a variety of factors. Prior research has highlighted the influence of personal [24, 29, 39, 51], social [3, 5, 16, 42], and cultural factors [29, 32, 34, 35, 41], foregrounding toxicity as a complex, multi-faceted phenomenon. In this section, our focus is on the relationship between toxicity and game design as another underexplored factor. Our resulting framework (see Figure 1) identifies three key elements through which

game design may facilitate toxicity: perceived triggers, perceived vulnerabilities, and affordance in game design. "Perceived triggers" refer to the aspects of game design that players believe can drive certain individuals into exhibiting toxic behavior. "Perceived vulnerabilities" refer to the aspects of game design that players believe can expose certain players to vulnerabilities, making them more susceptible to becoming targets of toxic actions. Affordance of the design refers to how the interactions between players and the game design result in unanticipated consequences and enable toxicity to occur.

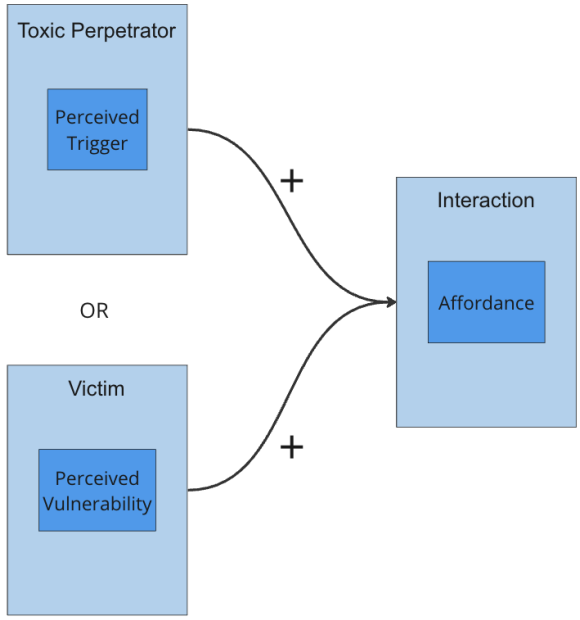


Figure 1. Framework of Toxicity by Design.

5.1.1 Perceived Trigger by Game Design

Game researchers observed that 'trigger,' which is defined as a factor that motivates individuals to engage in certain behaviors, can facilitate toxicity [37, 51]. Kordyaka et al. found that triggers can result from individual factors [37]. For example, a player may be driven by achievement and thus get triggered when losing [37]. Kou found that triggers can result from social factors [42]. For example, players can get triggered when they have resource conflicts with their teammates [42]. Our findings show not only how game design can become a trigger for toxicity but also how it affects individual factors and social factors and thus contributes to toxicity.

First, the game design that requires players' extensive time and effort influence individual factors and can be a trigger. Previous literature has highlighted how time is important for player experiences (PX) and once the PX is disturbed, players tend to be toxic. For example, Fahlström and Matson found that players engaged in longer game sessions tend to exhibit more toxicity [21]. This may be due to prolonged interactions with teammates, potentially leading to extended periods of frustration [42, 53]. Such frustration can build up, ultimately impacting players' emotional states and leading to toxic behavior [37]. Our findings further emphasize that not only time but also the efforts players invest during these gaming sessions are crucial factors influencing their emotional states. For example, in Gears, players invest significant time and

effort into progressing through waves. However, if they fail in a wave, their previous investment in time and effort is lost, triggering frustration and potentially leading to toxic behavior.

Second, certain game designs could contribute to the emergence of conflicts between teammates, influencing social factors and resulting in toxicity. Social factors describe how social dynamics within the team affect players' relationships with each other and result in toxicity. For example, players tend to attribute failure to other players [3]. Players may get into conflicts and even toxic conduct due to competing for in-game resources [42]. Particularly, Fahlström and Matson [21] observed that stealing kills, a valuable advantage by killing enemies, from teammates is deemed toxic as well. Our findings show certain designs, such as matchmaking systems can turn teammates into rivals, and thus affect social factors and facilitate toxicity. For instance, in League of Legends (LoL), where five distinct roles are available, it's possible for multiple team members to compete for the same role. This competition can lead to arguments and toxicity in team chats as members contest for their preferred roles. Additionally, when roles are assigned, the sight of someone else securing a coveted role can provoke toxic behavior, such as banning characters out of spite.

Lastly, ambiguity in game design, combined with individual factors, can influence social factors and trigger toxicity. Ambiguity in game design, such as ambiguity in one's supposed responsibility in teamwork, can lead to different playstyles. Previous studies have highlighted how strategic disagreements and gameplay conflicts can precipitate toxicity [21, 36, 42]. Deslauriers et al. mentioned how ambiguity in objective settings in Dead by Daylight can be a potential source of gameplay conflict [20]. Our findings extend this notion, revealing that not just objective ambiguity, but also the vagueness in role design and character design can lead to conflicts of playstyles and thus toxicity. For example, a player in a support role who has both healing and damage-dealing capabilities may face a challenge in balancing damage and healing, leading to conflicts over the appropriate extent of each among teammates' acceptance. Moreover, in games like LoL, the selection of 'roles' that define map positions versus the choice of character types for each role can result in a team composition imbalance, such as the absence of a tank, which in turn can spark blame and toxicity. Conversely, games like Overwatch assign roles based on character types, mitigating conflict over selection. In addition, individual factors, such as players' differing motivations [79] for engaging with a game in different playstyles, can facilitate conflicts and contribute to toxicity. In sum, players' motivations for choosing certain playstyles and interpretations of playstyle facilitate conflicts and thus toxicity when the game design is ambiguous.

5.1.2 Perceived Vulnerability by Game Design

Vulnerability refers to a state in which an individual is exposed to the potential risk of being targeted in a harmful action. Players who are vulnerable may face a heightened likelihood of being subjected to toxic behavior by others. Foo and Koivisto [24] have demonstrated that players predisposed to toxic behavior often seek out those who exhibit signs of vulnerability. To understand what could contribute to one's vulnerability, Kordyaka et al. discussed how an individual's personality traits may influence their susceptibility to being targeted [40]. Our research has revealed that certain aspects of game design can inadvertently place players in such a vulnerable state. This vulnerable state can then become a focal point for toxic individuals to exploit, even in the absence of direct 'triggers' for toxicity. We found that game design can

introduce three types of vulnerabilities that potentially exacerbate the risk of players becoming targets of toxicity.

The first type of vulnerability identified in our research is related to the visibility of data due to a lack of privacy protection. Kou and Gui [44] have described that players often scrutinize their teammates' performance data from previous gameplay to justify toxic behavior, highlighting the concerns regarding privacy issues associated with data visibility. Our findings reinforce the privacy concerns of game data, highlighting how publicized game data can expose players to toxic behavior. For instance, in-game performance data can display a player's performance over time in game, making it easy for others to initiate toxic behaviors upon observing poor performance. Game data related to promotions, referring to a critical phase where a player must pass successfully in order to advance to the next rank, can also make players who are at a promotional stage vulnerable to toxic actions, as toxic individuals may actively seek out such promotion information and intentionally disrupt those players' promotion games.

Another type of information that can bring vulnerability to players involves demographic information, such as gender. Previous scholars have highlighted that female players received sexual harassment targeted to their gender [25, 69]. In games with voice channels, female players' gender information can become apparent to all teammates using the same channel. While these channels are designed to facilitate game-related communication, they also inadvertently disclose gender information. While previous research [51] has indicated that the absence of voice channels might lead to a lack of social cues, weakening team bonds and potentially increasing toxicity, our findings indicate that the presence of voice channels can specifically expose female players and even other minorities to vulnerability, leading to instances of sexual harassment. Overall, the lack of privacy protection can leave certain groups vulnerable to becoming victims of toxic behavior.

The second type of vulnerability arises from aspects of game design that impact fairness. Previous research has explored various elements of fairness within game design, highlighting how certain practices negatively affect PX, such as 'pay-to-win' models [26], the ban-pick system and map design [78], and latency issues like ping delays [10]. However, there has been a shortfall in discussions on how elements of fairness in game design might contribute to toxicity or increase players' vulnerability to becoming victims. In our findings, we found that unfairness not only detracts PX but also places individuals in unfair situations, rendering them more susceptible to toxicity. For instance, in League of Legends (LoL), players with lower priority in the champion selection process may not be able to choose their preferred roles. Consequently, these players are at a disadvantage compared to those at the top of the selection queue, increasing their vulnerability. Similarly, a player who was matched with a four-person party in a five-player team might feel powerless during the voting process for kicking a player, as the design of the voting system and the option for four-player parties inherently disadvantages the solo player. These examples of unfair game design can unintentionally create situations where some players are rendered more vulnerable than others, thereby elevating their risk of encountering toxic behavior.

The third type of vulnerability arises from close collaboration with teammates, where players need to perform under pressure and, at the same time, are vulnerable to close scrutiny and abusive remarks from teammates. Fahlström and Matson [21] have shown that the significant influence of players' actions and their impact on team dynamics contribute to toxicity. Because of the ability to influence each other's and the team's overall performance, errors or inactions can both become grounds for blame. Minor errors can significantly affect the team, prompting

players to blame others and diminish their own responsibility in match outcomes. For instance, in our finding, one participant in Rainbow 6 highlighted the critical importance of each member's responsibility in a small team, which can exacerbate toxicity. Not only are mistakes a basis for blame, but inaction can also provoke toxicity. Liu and Agur [51] have indicated that within teamwork contexts, individuals are prone to minimizing their own accountability and attributing failures to their peers. The influence of inactions can be a reason to blame others for their accountability. For example, in our findings, a player blames Jungler for not ganking the lane. Players in certain roles, such as Jungler, bear more responsibility, leading to potentially unbalanced collaboration. Thus, the inability to collaborate with some players creates opportunities for toxic conduct, highlighting how tight teamwork can introduce vulnerabilities to players in certain roles. Overall, tight teamwork exposes players to vulnerabilities, stemming from individual mistakes that may not directly cause game failure or from role responsibilities that demand more collaboration than a player can effectively manage.

5.1.3 Affordance in Game Design

Affordance represents an interpretive relationship between users and a design, emerging as users engage with the design [75]. As players engage with a game, their interactions may extend beyond what designers initially envisioned. This phenomenon may resemble the concept of unintended consequences of design [52]; however, the distinction lies in that affordance specifically pertains to the nuances of design interaction, whereas unintended consequences encompass all outcomes, not solely those stemming from interaction, but also from other elements such as unfair scenarios which lead to unintended vulnerabilities of players.

We define affordance in game design as an interpretive relationship between the player and the interactive design of the game. The interactive design is the design that allows interaction between players. While the interactive design is innocuous, the affordance in interactive design shapes how players exploit these designs to facilitate toxic interactions with others. Previous scholars have found several interactive design features in games that can be abused by players to conduct toxicity. For example, Türkay and Adinolf [72] have demonstrated that players exploit the friend-adding feature for verbal abuse. Lam et al. [48] noted players use the emote feature for spamming others within the game. Our research verified previous literature by pointing out toxicity in similar interactive designs: players engage in verbal abuse during post-game chats and misuse dance features to spam others. We also found that, in first-person shooter games, team damage and vote-to-kick mechanisms can be an interactive design players can use to bully others. In games like League of Legends (LoL), the ban-pick system can be an interactive design that players can use to maliciously ban characters preferred by others. These instances we found highlighted how affordance in various interactive designs can inadvertently facilitate toxicity.

We see no affordance in game design when players do not have a direct interaction with interactive design. For example, players' personal data, although can be used in an unintended way for toxic conduct, is not an interactive design that allows interactions between players. Ambiguous character and role designs, although can be played in various of ways out of game designers' expectations or teammates' expectations, are not interactive designs since it is only about the interaction between players and in-game avatars, instead of other players.

5.1.4 Summary

In this section, we discussed how game design can trigger toxicity, either by fostering behaviors that lead to toxic actions or by creating vulnerabilities that make players targets of such behavior. While there may be other non-design-related triggers and vulnerabilities, our focus is on the impact of game design on both perpetrators and victims of toxicity. We also propose affordances in interactive design that allow toxic interaction to happen in games.

5.2 Design Implication

While game designs can inadvertently contribute to the emergence of toxicity, eliminating these design elements entirely is not the optimal or realistic solution for reducing such behavior. Plus, game design is not the sole factor that leads to toxicity. For instance, the option to remove a player from a game in CS:GO, though potentially misused by larger groups to target individuals unfairly, serves as a vital tool for addressing genuinely toxic behavior. Without this mechanism, alternative issues may arise. This illustrates the complexity of balancing game design to mitigate toxicity without compromising the game's inherent dynamics and player experience. In this section, we critically engage with the game design that leads to different levels of toxicity as well as the game designs that participants envisioned to mitigate toxicity.

It is important to underscore that the findings of our study, while derived from the games participants played and mentioned, hold relevance for a broader spectrum of online gaming experiences. For instance, the concerns regarding privacy expressed by female players in the context of an online board game featuring a voice channel may extend to other online gaming environments with similar communication features. Moreover, the design implications below are not confined solely to the specific games we mentioned; rather, they bear applicability to a wider range of online games sharing similar mechanisms.

5.2.1 Managing Triggers by Design

We encourage game designers to proactively consider the triggers present within the context of specific online games and develop strategies to mitigate these triggers, thereby addressing toxicity. To provide more tangible design implications, we present two design suggestions based on triggers identified in this study. One suggestion involves providing guidance to assist players in better understanding character roles and team balance by the calculated win rate. This design strategy may help players grasp the balance of character types within the team and avoid potential toxicity resulting from character selection. However, there's a possibility that such a win rate might not accurately capture the impact of individual character selections on overall success. Additionally, games could integrate visualized data showcasing character balance, such as total defense capabilities or damage output, during the character selection stage. For instance, in Mario Kart, the impact of Kart selection on performance is made visible in the Kart selection stage [83]. This visualized data may not only allow players to see the effect of their character choices on various aspects of team composition but also give players an overview of team balance. Therefore, it may deal with the situation where players make selections that result in an imbalanced team composition.

In addition, game designers could enhance support for players facing ambiguity in their playstyles during the game. As discussed earlier, the affordance in gameplay can lead to various interpretations by players, which leads to different playstyles that can trigger conflicts and toxicity. Providing in-game advice for players tasked with ambiguous characters could offer clarity on their responsibilities and help align their actions with team expectations. For instance, a support character possessing both damage-dealing and healing abilities might be uncertain

about which action best meets their teammates' expectations and benefits the team. By integrating in-game statistics that display the team's overall damage output and healing requirements, the game design could aid support players in making more informed decisions about their in-game actions based on the team's needs for damage and healing, thus clarifying their role in real-time. This strategy could potentially reduce the triggers of conflict that arise from differing gameplay strategies.

5.2.2 Managing Vulnerabilities by Design

We encourage game designers to proactively consider the vulnerabilities present within the context of specific online games and develop strategies to support potential victims, thereby addressing toxicity. To provide more tangible design implications, we present two design suggestions based on the vulnerabilities we identified in this study. Our first recommendation is to reevaluate the privacy enhancement of player data. This includes in-game performance statistics, account information, and potentially sensitive data exposed through voice channel affordances, such as gender. Game companies could choose to empower players with the ability to manage the privacy of their data, resonating with prior research advocating for player control over aspects like voice channel volume [1]. Regarding the information disclosed through voice channels, game companies might contemplate implementing voice transformation technology to conceal individuals' identities by altering their voices in real-time. This strategy could also empower players to control their personal information visibility by modifying their voice. With the control of the data visibility, players who believe they are in a vulnerable state can protect themselves by disabling the visibility of the data.

Another design implication we suggest is to enhance players' fairness perception in PX. Games designs that could render an unfair situation should be reevaluated and redesigned. For instance, LoL has changed the system for choosing roles, which assigns players roles before they enter a match. With such a system, players do not need to negotiate the role assignment in the chat and thus reduce the vulnerability of players in the later of the sequence to choose the role, therefore mitigating toxicity. To address the unfairness encountered in open-world games, game companies might consider regulating certain encounters or interactions. For instance, implementing restrictions on the lowest-level areas that players can enter could prevent higher-level players from accessing areas intended for lower-level players. Alternatively, game companies could adjust the levels of non-player characters (NPCs) to match the individual levels of players, ensuring that all players face equal difficulty in acquiring specific resources from NPCs. By either of the strategies, players with lower levels may have a better chance to obtain resources and enjoy a fairer gaming environment.

5.2.3 In-time Strategies to Maximize Player Experiences in Game

Online games have different mechanisms that can bring about various unintended consequences. We realize that it is almost impossible to list all the unintended consequences for all the games we mentioned and all types of online games. However, what we can suggest is that game designers can come up with in-time strategies according to the specific game context to maximize player experiences in the game to minimize the consequences of toxicity. For example, we observe that with certain game designs, what counts as 'toxic behavior' is different. AFK is seen as toxic behavior in many video games such as LoL and Overwatch, but less so in video games like Battlefield. This is because there are always new players to replace the AFK players and thus the player experiences are not affected by the AFK players. While it may not be

feasible to always find substitute players in every game type, developers could explore in-time strategies, such as deploying AI teammates to fill gaps left by AFK players [76, 81] to support player experiences in the game.

While these in-time strategies may not prevent the occurrence of toxicity, mitigating its impact can enhance player experiences. However, game designers have to be cautious about implementing such in-time features to prevent exploitation by toxic players, such as manipulating AI to unduly influence game outcomes. With these in-time strategies in game design, we believe that when the impact of toxicity is minimized to the extent that player experiences are not significantly affected, the behavior may not even be considered toxic.

6 LIMITATIONS AND FUTURE WORK

In this section, we discuss the limitations of our study, as well as what remains to be done in the future. First, our methodology, which involved interviewing players, means our findings primarily reflect how players perceive the influence of game design on what they identify as toxicity. It's important to acknowledge that these perceptions might be biased by their understanding of game design and personal experiences with toxicity. More research can be done using different methodologies (e.g. collecting and analyzing player gameplay data from APIs) to measure and predict the relationship between toxicity and game design factors.

Second, during our interview sessions, we did not specifically inquire about participants' self-assessed expertise, nor did we collect data on the duration or frequency of their gameplay for each game. We recognize that participants' personal experiences can offer varied contexts and explanations regarding the occurrence of toxicity in detail. Future researchers could explore this aspect using alternative methodologies for data analysis, such as quantitative studies, to assess how personal experiences influence perspectives on toxicity in game design.

Third, this paper mainly focuses on identifying the themes that cause toxicity in online games played by our participants using a qualitative methodology. Given the nature of qualitative data and the specific games our participants have played, our interview data may not cover all types of toxicity across all online games. Therefore, it is possible that some of the toxicity caused by some specific game designs is not reported in this paper. More research can be done using large-scale data (e.g. a survey) to validate the game design factors we identified in this study. In addition, future researchers can also look into how game design causes toxicity in different games and whether they have differences from each other to cross-validate our study's result.

Lastly, this study initiates a fresh inquiry into the nexus between toxicity and game design through the lens of players' perspectives. It is important to note that our participants consist solely of university students, and thus their interpretations may not fully encapsulate the viewpoints of players across diverse demographic cohorts. Hence, further research is needed to integrate the perspectives of other stakeholders, such as game designers, and individuals from varied demographic groups, in order to comprehensively examine how game design impacts toxicity.

7 CONCLUSION

In this study, we investigated players' perceptions of how game design influences toxicity. Our research offers insights into factors contributing to toxicity across various online games, identifying four game design patterns that exacerbate toxicity and affect its prevalence: team

interdependency, fairness, interactions, and privacy. We argue that toxicity may be viewed as an unintended consequence of game design, leading to the triggers of toxic perpetrators, increased vulnerability among victims, and specific affordances in player interactions with the design. Moving forward, it is important for game designers to assess the potential impact of the identified game design factors on toxicity. This research enhances our understanding of toxicity in online games through a design-oriented lens from players' perspectives. It sheds light on how toxicity arises as a type of unintended consequence alongside particular game design patterns. Player experiences with toxicity by game design underscore the importance of prosocial design in online games. With the hope of mitigating toxicity and encouraging prosocial activities, we propose design implications based on the toxic by design framework to guide game designers in managing triggers and vulnerabilities caused by game design, and dynamically enhancing player experiences during gameplay. This work also aims to bridge the gap between industry practices and player perspectives by incorporating players' views into game design.

ACKNOWLEDGMENTS

We truly appreciate all the anonymous reviewers at CHI PLAY 2024 and all lab members in Health and Play Lab at Penn State for their insightful and constructive feedback. This work was partially supported by National Science Foundation under grant #2326505.

REFERENCES

[1] Lucy A. Sparrow, Martin Gibbs, and Michael Arnold. 2021. The ethics of multiplayer game design and community management: industry perspectives and challenges. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021. 1–13.

[2] Ernest Adams and Joris Dormans. 2012. *Game Mechanics: Advanced Game Design*. New Riders.

[3] Sonam Adinolf and Selen Turkay. 2018. Toxic Behaviors in Esports Games: Player Perceptions and Coping Strategies. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*, October 23, 2018. ACM, Melbourne VIC Australia, 365–372. https://doi.org/10.1145/3270316.3271545

[4] Joan S. Ash, Marc Berg, and Enrico Coiera. 2004. Some Unintended Consequences of Information Technology in Health Care: The Nature of Patient Care Information System-related Errors. *Journal of the American Medical Informatics Association* 11, 2 (March 2004), 104–112. https://doi.org/10.1197/jamia.M1471

[5] Jane Barnett, Mark Coulson, and Nigel Foreman. 2010. Examining Player Anger in World of Warcraft. In *Online Worlds: Convergence of the Real and the Virtual*, William Sims Bainbridge (ed.). Springer London, London, 147–160. https://doi.org/10.1007/978-1-84882-825-4_12

[6] Nicole A Beres, Julian Frommel, Elizabeth Reid, Regan L Mandryk, and Madison Klarkowski. 2021. Don't You Know That You're Toxic: Normalization of Toxicity in Online Gaming. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, May 06, 2021. ACM, Yokohama Japan, 1–15. https://doi.org/10.1145/3411764.3445157

[7] Virginia Braun and Victoria Clarke. 2021. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qualitative Research in Sport, Exercise and Health* 13, 2 (March 2021), 201–216. https://doi.org/10.1080/2159676X.2019.1704846

[8] Virginia Braun and Victoria Clarke. 2022. Conceptual and design thinking for thematic analysis. *Qualitative psychology* 9, 1 (2022), 3.

[9] Virginia Braun, Victoria Clarke, Nikki Hayfield, Louise Davey, and Elizabeth Jenkinson. 2022. Doing Reflexive Thematic Analysis. In *Supporting Research in Counselling and Psychotherapy : Qualitative, Quantitative, and Mixed Methods Research*, Sofie Bager-Charleson and Alistair McBeath (eds.). Springer International Publishing, Cham, 19–38. https://doi.org/10.1007/978-3-031-13942-0_2

[10] Jeremy Brun, Farzad Safaei, and Paul Boustead. 2006. Fairness and playability in online multiplayer games. (2006). Retrieved November 27, 2023 from https://ro.uow.edu.au/infopapers/232/

[11] Jonathan Burnay, Brad J. Bushman, and Frank Larøi. 2019. Effects of sexualized video games on online sexual harassment. *Aggressive Behavior* 45, 2 (March 2019), 214–223. https://doi.org/10.1002/ab.21811

[12] Kelly Caine. 2016. Local Standards for Sample Size at CHI. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (*CHI '16*), May 07, 2016. Association for Computing Machinery, New York, NY, USA, 981–992. https://doi.org/10.1145/2858036.2858498

[13] Ann Frances Cameron and Jane Webster. 2005. Unintended consequences of emerging communication technologies: Instant Messaging in the workplace. *Computers in Human Behavior* 21, 1 (January 2005), 85–103. https://doi.org/10.1016/j.chb.2003.12.001

[14] Vivian Hsueh-Hua Chen, Henry Been-Lirn Duh, and Chiew Woon Ng. 2009. Players who play to make others cry: the influence of anonymity and immersion. In *Proceedings of the International Conference on Advances in Computer Entertainment Technology*, October 29, 2009. ACM, Athens Greece, 341–344. https://doi.org/10.1145/1690388.1690454

[15] Victoria Clarke and Virginia Braun. 2021. *Thematic Analysis : A Practical Guide*. SAGE Publications Ltd, London. Retrieved from http://digital.casalini.it/9781526417305

[16] Christine Cook, Rianne Conijn, Juliette Schaafsma, and Marjolijn Antheunis. 2019. For whom the gamer trolls: A study of trolling interactions in the online gaming context. *Journal of Computer-Mediated Communication* 24, 6 (2019), 293–318.

[17] Christine Dearnley. 2005. A reflection on the use of semi-structured interviews. *Nurse Researcher (through 2013)* 13, 1 (2005), 19–28.

[18] Martin Delhove and Tobias Greitemeyer. 2020. The relationship between video game character preferences and aggressive and prosocial personality traits. *Psychology of popular media* 9, 1 (2020), 96.

[19] Ansgar E. Depping, Colby Johanson, and Regan L. Mandryk. 2018. Designing for Friendship: Modeling Properties of Play, In-Game Social Capital, and Psychological Well-being. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*, October 23, 2018. ACM, Melbourne VIC Australia, 87–100. https://doi.org/10.1145/3242671.3242702

[20] Patrick Deslauriers, Laura Iseut Lafrance St-Martin, and Maude Bonenfant. 2020. Assessing toxic behaviour in dead by daylight: perceptions and factors of toxicity according to the game's official subreddit contributors. *Game Studies* 20, 4 (2020). Retrieved November 10, 2023 from https://www.researchgate.net/profile/Laura-Iseut-Lafrance-St-Martin/publication/353609119_Assessing_Toxic_Behaviour_in_Dead_by_Daylight_Perceptions_and_Factors_of_Toxicity_According_to_the_Game's_Official_Subreddit_Contributors/links/6105963a0c2bfa282a12ad5b/Assessing-Toxic-Behaviour-in-Dead-by-Daylight-Perceptions-and-Factors-of-Toxicity-According-to-the-Games-Official-Subreddit-Contributors.pdf

[21] Josefine Fahlström and Emma Matson. 2014. *Preventing Toxic Behaviour through Game Mechanics*. Retrieved November 30, 2023 from https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-227805

[22] Mary Flanagan, Daniel C. Howe, and Helen Nissenbaum. 2005. Values at play: Design tradeoffs in socially-oriented game design. In *Proceedings of the SIGCHI conference on human factors in computing systems*, 2005. 751–760.

[23] Mary Flanagan and Helen Nissenbaum. 2007. A game design methodology to incorporate social activist themes. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2007. 181–190.

[24] Chek Yang Foo and Elina M. I. Koivisto. 2004. Defining grief play in MMORPGs: player and developer perceptions. In *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology*, September 02, 2004. ACM, Singapore, 245–250. https://doi.org/10.1145/1067343.1067375

[25] Jesse Fox and Wai Yen Tang. 2017. Women's experiences with general and sexual harassment in online video games: Rumination, organizational responsiveness, withdrawal, and coping strategies. *New Media & Society* 19, 8 (August 2017), 1290–1307. https://doi.org/10.1177/1461444816635778

[26] Guo Freeman, Karen Wu, Nicholas Nower, and Donghee Yvette Wohn. 2022. Pay to Win or Pay to Cheat: How Players of Competitive Online Games Perceive Fairness of In-Game Purchases. *Proc. ACM Hum.-Comput. Interact.* 6, CHI PLAY (October 2022), 1–24. https://doi.org/10.1145/3549510

[27] John M. Freeman. 2007. Beware: The misuse of technology and the law of unintended consequences. *Neurotherapeutics* 4, 3 (July 2007), 549–554. https://doi.org/10.1016/j.nurt.2007.04.003

[28] Enrico Gandolfi. 2018. You have got a (different) friend in me: Asymmetrical roles in gaming as potential ambassadors of computational and cooperative thinking. *E-Learning and Digital Media* 15, 3 (May 2018), 128–145. https://doi.org/10.1177/2042753018757757

[29] Enrico Gandolfi, Richard E. Ferdig, Karlie Krause, Amy Copus, Sarah Ostrowski-Delahanty, and Sonia Alemagno. 2023. Problematic Gaming at a Crossroad: Exploring the Interplay Between Internet Gaming Disorder, Toxic Attitudes, and Empathy in Digital Entertainment. *Games and Culture* (November 2023), 15554120231211991. https://doi.org/10.1177/15554120231211991

[30] Colin M. Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L. Toombs. 2018. The Dark (Patterns) Side of UX Design. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, April 21, 2018. ACM, Montreal QC Canada, 1–14. https://doi.org/10.1145/3173574.3174108

[31] Olli I. Heimo, J. Tuomas Harviainen, Kai K. Kimppa, and Tuomas Mäkilä. 2018. Virtual to Virtuous Money: A Virtue Ethics Perspective on Video Game Business Logic. *J Bus Ethics* 153, 1 (November 2018), 95–103. https://doi.org/10.1007/s10551-016-3408-z

[32] Zorah Hilvert-Bruce and James T. Neill. 2020. I'm just trolling: The role of normative beliefs in aggressive behaviour in online gaming. *Computers in Human Behavior* 102, (2020), 303–311.

[33] Kasper Hornbæk and Antti Oulasvirta. 2017. What Is Interaction? In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, May 02, 2017. ACM, Denver Colorado USA, 5040–5052. https://doi.org/10.1145/3025453.3025765

[34] Bastian Kordyaka, Katharina Jahn, and Bjoern Niehaves. 2020. Towards a unified theory of toxic behavior in video games. *Internet Research* 30, 4 (2020), 1081–1102.

[35] Bastian Kordyaka, Jeanine Krath, Solip Park, Henrik Wesseloh, and Samuli Laato. 2022. Understanding toxicity in multiplayer online games: The roles of national culture and demographic variables. 2022. Hawaii International Conference on System Sciences. Retrieved November 30, 2023 from https://trepo.tuni.fi/handle/10024/145969

[36] Bastian Kordyaka and Björn Kruse. 2021. Curing toxicity–developing design principles to buffer toxic behaviour in massive multiplayer online games. *Safer communities* 20, 3 (2021), 133–149.

[37] Bastian Kordyaka, Samuli Laato, Juho Hamari, Tobias Scholz, and Björn Niehaves. 2023. What drives gamer toxicity? Essays from players. (April 2023). Retrieved November 30, 2023 from https://trepo.tuni.fi/handle/10024/151587

[38] Bastian Kordyaka, Samuli Laato, Katharina Jahn, Juho Hamari, and Bjoern Niehaves. 2023. The Cycle of Toxicity: Exploring Relationships between Personality and Player Roles in Toxic Behavior in Multiplayer Online Battle Arena Games. *Proc. ACM Hum.-Comput. Interact.* 7, CHI PLAY (October 2023), 397:611-397:641. https://doi.org/10.1145/3611043

[39] Bastian Kordyaka, Samuli Laato, Aida Stelter, and Bjoern Niehaves. 2023. Dark Desires? Using the Theory of Basic Desires to Better Understand Toxic Behavior in Multiplayer Online Games. In *Proceedings of the 56th Annual Hawaii International Conference on System Sciences, HICSS 2023*, 2023. 5551–5559. Retrieved November 30, 2023 from https://researchportal.tuni.fi/en/publications/dark-desires-using-the-theory-of-basic-desires-to-better-understa

[40] Bastian Kordyaka, Samuli Laato, Sebastian Weber, and Bjoern Niehaves. 2023. What constitutes victims of toxicity - identifying drivers of toxic victimhood in multiplayer online battle arena games. *Frontiers in Psychology* 14, (2023). Retrieved November 30, 2023 from https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1193172

[41] Bastian Kordyaka, Solip Park, Jeanine Krath, and Samuli Laato. 2023. Exploring the Relationship Between Offline Cultural Environments and Toxic Behavior Tendencies in Multiplayer Online Games. *Trans. Soc. Comput.* 6, 1–2 (June 2023), 3:1-3:20. https://doi.org/10.1145/3580346

[42] Yubo Kou. 2020. Toxic Behaviors in Team-Based Competitive Gaming: The Case of League of Legends. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, November 02, 2020. ACM, Virtual Event Canada, 81–92. https://doi.org/10.1145/3410404.3414243

[43] Yubo Kou and Xinning Gui. 2014. Playing with strangers: understanding temporary teams in league of legends. In *Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play*, October 19, 2014. ACM, Toronto Ontario Canada, 161–169. https://doi.org/10.1145/2658537.2658538

[44] Yubo Kou and Xinning Gui. 2018. Entangled with Numbers: Quantified Self and Others in a Team-Based Online Game. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW (November 2018), 1–25. https://doi.org/10.1145/3274362

[45] Yubo Kou, Renkai Ma, Zinan Zhang, Yingfan Zhou, and Xinning Gui. 2024. Community Begins Where Moderation Ends: Peer Support and Its Implications for Community-Based Rehabilitation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (*CHI '24*), May 11, 2024. Association for Computing Machinery, New York, NY, USA, 1–18. https://doi.org/10.1145/3613904.3642675

[46] Rachel Kowert and Chrissy Cook. 2022. The toxicity of our (virtual) cities: prevalence of dark participation in games and perceived effectiveness of reporting tools. (2022). Retrieved February 4, 2024 from https://scholarspace.manoa.hawaii.edu/handle/10125/79724

[47] Haewoon Kwak and Jeremy Blackburn. 2015. Linguistic Analysis of Toxic Behavior in an Online Video Game. In *Social Informatics* (*Lecture Notes in Computer Science*), 2015. Springer International Publishing, Cham, 209–217. https://doi.org/10.1007/978-3-319-15168-7_26

[48] Lawrence T. Lam, ZaoHuo Cheng, and XinMin Liu. 2013. Violent Online Games Exposure and Cyberbullying/Victimization Among Adolescents. *Cyberpsychology, Behavior, and Social Networking* 16, 3 (March 2013), 159–165. https://doi.org/10.1089/cyber.2012.0087

[49] Matthew Lapolla. 2020. Tackling Toxicity: Identifying and Addressing Toxic Behavior in Online Video Games. Seton Hall University. Retrieved from https://scholarship.shu.edu/dissertations/2798

[50] Sung Je Lee, Eui Jun Jeong, and Joon Hyun Jeon. 2019. Disruptive behaviors in online games: Effects of moral positioning, competitive motivation, and aggression in "League of Legends." *Social Behavior and Personality* 47, 2 (2019), 1–9. https://doi.org/10.2224/sbp.7570

[51] Yansheng Liu and Colin Agur. 2023. "After All, They Don't Know Me" Exploring the Psychological Mechanisms of Toxic Behavior in Online Games. *Games and Culture* 18, 5 (July 2023), 598–621. https://doi.org/10.1177/15554120221115397

[52] Laurie McAulay. 2007. Unintended consequences of computer-mediated communications. *Behaviour & Information Technology* 26, 5 (September 2007), 385–398. https://doi.org/10.1080/01449290500535343

[53] Dave McLean, Frank Waddell, and James Ivory. 2020. Toxic teammates or obscene opponents? Influences of cooperation and competition on hostility between teammates and opponents in an online game. *Journal For Virtual Worlds Research* 13, 1 (2020). Retrieved November 10, 2023 from https://jvwr-ojs-utexas.tdl.org/jvwr/article/view/7334

[54] Luke Munn. 2023. Toxic play: Examining the issue of hate within gaming. *First Monday* (September 2023). https://doi.org/10.5210/fm.v28i9.12508
[55] Mahin Naderifar, Hamideh Goli, and Fereshteh Ghaljaie. 2017. Snowball Sampling: A Purposeful Method of Sampling in Qualitative Research. *Strides in Development of Medical Education* 14, 3 (September 2017). https://doi.org/10.5812/sdme.67670
[56] Linda L. Naimi and Richard Mark French. 2010. The unintended consequences of technological innovation: Bluetooth technology and cultural change. *The IPSI BgD Transactions on Internet Research* 3, (2010). Retrieved February 16, 2024 from https://www.researchgate.net/profile/Jelena-Zascerinska/publication/215969597_E-Collaboration_Technologies_in_TeachingLearning_Activity/links/09e4150b8b0433fbaa000000/E-Collaboration-Technologies-in-Teaching-Learning-Activity.pdf#page=4
[57] Joaquim A. M. Neto, Kazuki M. Yokoyama, and Karin Becker. 2017. Studying toxic behavior influence and player chat in an online video game. In *Proceedings of the International Conference on Web Intelligence*, August 23, 2017. ACM, Leipzig Germany, 26–33. https://doi.org/10.1145/3106426.3106452
[58] Helen Nissenbaum. 2004. Privacy as Contextual Integrity Symposium: Technology, Values, and the Justice System. *Wash. L. Rev.* 79, 1 (2004), 119–158.
[59] Efthymia Paparistodemou, Richard Noss, and Dave Pratt. 2008. The Interplay Between Fairness and Randomness in a Spatial Computer Game. *Int J Comput Math Learning* 13, 2 (July 2008), 89–110. https://doi.org/10.1007/s10758-008-9132-8
[60] C. Parker, S. Scott, and A. Geddes. 2019. Snowball Sampling. *SAGE Research Methods Foundations* (September 2019). Retrieved November 8, 2023 from http://methods.sagepub.com/foundations/snowball-sampling
[61] Michael Quinn Patton. 1990. *Qualitative evaluation and research methods*. Sage Publications, Inc, Newbury Park, CA.
[62] Christopher A Paul. 2018. *The toxic meritocracy of video games: Why gaming culture is the worst*. U of Minnesota Press.
[63] Elizabeth Reid, Regan L. Mandryk, Nicole A. Beres, Madison Klarkowski, and Julian Frommel. 2022. Feeling Good and In Control: In-game Tools to Support Targets of Toxicity. *Proceedings of the ACM on Human-Computer Interaction* 6, CHI PLAY (2022), 1–27.
[64] Elizabeth Reid, Regan L. Mandryk, Nicole A. Beres, Madison Klarkowski, and Julian Frommel. 2022. "Bad Vibrations": Sensing Toxicity From In-Game Audio Features. *IEEE Transactions on Games* 14, 4 (December 2022), 558–568. https://doi.org/10.1109/TG.2022.3176849
[65] Bradford W. Reyns, Melissa W. Burek, Billy Henson, and Bonnie S. Fisher. 2013. The unintended consequences of digital technology: exploring the relationship between sexting and cybervictimization. *Journal of Crime and Justice* 36, 1 (March 2013), 1–17. https://doi.org/10.1080/0735648X.2011.641816
[66] Victoria L. Rubin and Sarah C. Camm. 2013. Deception in video games: Examining varieties of griefing. *Online Information Review* 37, 3 (2013), 369–387.
[67] Douglas Schuler and Aki Namioka. 1993. *Participatory design: Principles and practices*. CRC Press.
[68] Clay Spinuzzi. 2005. The Methodology of Participatory Design. *Technical Communication* 52, 2 (May 2005), 163–174.
[69] Wai Yen Tang and Jesse Fox. 2016. Men's harassment behavior in online video games: Personality traits and game factors. *Aggressive Behavior* 42, 6 (November 2016), 513–521. https://doi.org/10.1002/ab.21646
[70] Andrew A. Tawfik, Todd D. Reeves, and Amy Stich. 2016. Intended and Unintended Consequences of Educational Technology on Social Inequality. *TechTrends* 60, 6 (November 2016), 598–605. https://doi.org/10.1007/s11528-016-0109-5
[71] Edward Tenner. 1997. *Why Things Bite Back: Technology and the Revenge of Unintended Consequences*. Knopf Doubleday Publishing Group.
[72] Selen Türkay and Sonam Adinolf. 2019. Friending to Flame: How Social Features Affect Player Behaviours in an Online Collectible Card Game. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, May 02, 2019. ACM, Glasgow Scotland Uk, 1–12. https://doi.org/10.1145/3290605.3300567
[73] Selen Türkay, Jessica Formosa, Sonam Adinolf, Robert Cuthbert, and Roger Altizer. 2020. See no evil, hear no evil, speak no evil: How collegiate players define, experience and cope with toxicity. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020. 1–13.
[74] Andrew Turner, Rebecca Morris, Dylan Rakhra, Fiona Stevenson, Lorraine McDonagh, Fiona Hamilton, Helen Atherton, Michelle Farr, Sarah Blake, Jon Banks, Gemma Lasseter, Sue Ziebland, Emma Hyde, John Powell, and Jeremy Horwood. 2022. Unintended consequences of online consultations: a qualitative study in UK primary care. *Br J Gen Pract* 72, 715 (February 2022), e128–e137. https://doi.org/10.3399/BJGP.2021.0426
[75] Dhaval Vyas, Cristina M. Chisalita, and Gerrit C. van der Veer. 2006. Affordance in interaction. In *Proceedings of the 13th Eurpoean conference on Cognitive ergonomics: trust and control in complex socio-technical systems* (*ECCE '06*), September 20, 2006. Association for Computing Machinery, New York, NY, USA, 92–99. https://doi.org/10.1145/1274892.1274907
[76] Hua Wei, Jingxiao Chen, Xiyang Ji, Hongyang Qin, Minwen Deng, Siqin Li, Liang Wang, Weinan Zhang, Yong Yu, and Liu Linc. 2022. Honor of kings arena: an environment for generalization in competitive reinforcement learning. *Advances in Neural Information Processing Systems* 35, (2022), 11881–11892.

[77] Minerva Wu, Je Seok Lee, and Constance Steinkuehler. 2021. Understanding tilt in esports: A study on young league of legends players. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021. 1–9.
[78] Mingyang Wu, Shuo Xiong, and Hiroyuki Iida. 2016. Fairness mechanism in multiplayer online battle arena games. In *2016 3rd International Conference on Systems and Informatics (ICSAI)*, 2016. IEEE, 387–392.
[79] Nick Yee. 2006. Motivations for play in online games. *CyberPsychology & behavior* 9, 6 (2006), 772–775.
[80] José P. Zagal, Staffan Björk, and Chris Lewis. 2013. Dark patterns in the design of games. In *Foundations of Digital Games 2013*, 2013. .
[81] Rui Zhang, Nathan J. McNeese, Guo Freeman, and Geoff Musick. 2021. " An ideal human" expectations of AI teammates in human-AI teaming. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (2021), 1–25.
[82] Ágnes Zsila, Reza Shabahang, Mara S. Aruguete, and Gábor Orosz. 2022. Toxic behaviors in online multiplayer games: Prevalence, perception, risk factors of victimization, and psychological consequences. *Aggressive Behavior* 48, 3 (May 2022), 356–364. https://doi.org/10.1002/ab.22023
[83] 2017. *Mario Kart 8 Deluxe All Vehicle Parts*. Retrieved February 17, 2024 from https://www.youtube.com/watch?v=bDfHmUfePBk
[84] Free to Play? Hate, Harassment, and Positive Social Experiences in Online Games | ADL. Retrieved November 10, 2023 from https://www.adl.org/resources/report/free-play-hate-harassment-and-positive-social-experiences-online-games
[85] Hate is No Game: Harassment and Positive Social Experiences in Online Games 2021 | ADL. Retrieved January 15, 2024 from https://www.adl.org/resources/report/hate-no-game-harassment-and-positive-social-experiences-online-games-2021
[86] Hate Is No Game: Hate and Harassment in Online Games 2022 | ADL. Retrieved February 4, 2024 from https://www.adl.org/resources/report/hate-no-game-hate-and-harassment-online-games-2022

A APPENDICES

A.1 Interview Questions

1. Demographic questions - gender, race, occupation, location (State)
2. General Questions regarding experiences with online games
    a. What online games do you play? How long have you played?
    b. Whom do you usually play with?
    c. What is a typical scenario where you start playing an online game? (Big context)
3. Players' experiences with toxicity
    a. Have you encountered toxicity in online games before? Can you give an example?
    b. What games/type of games have you experienced toxicity interactions? How often is it?
    c. When do you see toxicity in a game? Begin? In the middle? Something happened?
    d. Have you ever initiated toxic behavior in a game? What's the reason for that?
    e. What type of toxicity interactions have you experienced? Chat/in game behavior?
4. Players' explanations of toxicity
    a. How would you define what is or is not toxicity? And why?
    b. What do you think causes toxicity among players?
    c. Player themselves?
    d. Do you think the game design leads to toxicity?
    e. Are there games that are more toxic than other games? And why? Any examples?

    f. Do you think toxicity is a problem in online games? And why?
5. Players' actions related to toxicity
    a. What do you usually do when you are the victim of toxic conduct?
    b. What do you usually do when you witness players being toxic to other people?
    c. How are you affected by toxicity in the game (long term)? Does it change how you behave/play the game?
    d. Do you think game companies have done enough to manage toxicity? And why?
    e. What else do you wish online games have to manage toxicity?
6. Ending questions
    a. Do you have anything else to add regarding your experiences with toxicity?

Received February 2024; revised June 2024; accepted July 2024.