

UPPSALA
UNIVERSITET

How does toxicity change depending on rank in League of Legends?

Faculty of
Department of Game Design

Author(s): Edward Leiman, William Herner

Bachelor's Thesis in Game Design, 15 hp

Program: Game Design Project Management

Supervisor(s): Magnus Johansson, Joshua Juvrud

Examiner: Michael Hiessboeck

May, 2019

Abstract
This thesis aims to investigate toxic remarks in three different ranks in League of Legends, Bronze, Gold, and Diamond. The purpose is to understand how toxic communication between players would change depending on rank. A framework from Neto, Alvino and Becker (2018) was adopted to define and count toxic remarks. The method relied on participant observation to gather data, three different ranks were specified for data collection. Fifteen games were played in each of the ranks; Bronze, Gold, and Diamond. Each game was recorded, transcribed and analyzed by dividing each toxic remark registered into Neto and Beckers predetermined categories. The study concluded that domain language is more often used by players with a higher rank, meaning that high ranked players tend to use toxicity that requires previous game knowledge to understand. On the contrary, low ranked players tend to stick to basic complaints and insults when using toxicity to remark teammates while playing.

Key words: Toxic Remark, Toxicity, Domain Language, Domain Knowledge, Spam,

Den här forskningen ämnar till att undersöka toxiska kommentarer i tre olika divisioner i League of Legends, Brons, Guld, och Diamant. Syftet är att försöka förstå hur toxiska kommentarer ändras mellan olika spelare i de olika divisionerna. Ett ramverk utformat av Neto, Alvino och Becker (2018) har använts för att kunna definiera samt räkna toxiska kommentarer. Metoden för datainsamlingen från de tre olika divisionerna genomfördes via deltagarobservation. Femton matcher spelades i varje division; Brons, Guld, och Diamant. Varje match spelades in, transkriberades, och analyserades genom att dela in toxiska kommentarer i Neto och Beckers olika kategorier. Forskningen drar slutsatsen att domän språk är oftare använt av spelare i högre divisioner, domän språk är kopplat till slang inom spel som kräver tidigare kunskap i spelet för att förstå. I kontrast till detta så använder spelare i lägre divisioner mer basala klagomål och förolämpningar när toxiska kommentarer riktas mot andra spelare.

Nyckelord:

Table of Contents

1 Background……..……5

2 Introduction……..…....…6

2.1 MOBA..…...……..…....…6

3 Previous works within the subject area.....….....7

4 Definition of Toxicity…..……..……....8

4.1 Toxic remark……..….....…6

4.2 Research Question……..……..…....…6

5 Method…….……...……...…....…9

5.1 Toxic Remarks.....9

5.1.1 Insults…,,,…...…….....…11

5.1.2 Taunts……....…….11

5.1.3 Complaints…….……..…12

5.1.4 Spam……...……..…....…12

5.1.5 Domain Language……..……..…....…12

5.2 Participant Observation……..……..…....…6

5.3 Code of Conduct.……..…....…6

5.4 Limitations..……..…....…6

6 Results……..…….……..……...……13

7 Analysis…….....14

8.1 The Ping System……..…….……...……...…...…..14

7.2 Section 2.....16

7.3 Section 3……..…….……...……...…...…..16

7.4 Section 4……..…….……...……...…...…..18

8 Conclusion…….....19

References

Appendix

1 Background

The increasing enjoyment of competitive games has also led to the growing concern of negative behavior, such as for example, cyberbullying, griefing, mischief and cheating. These behaviors can be grouped together under toxic behaviors. Due to differences in cultures, customs, ethics and subjective perceptions of griefing across games and genres players doing the griefing are not even sometimes aware of it themselves (Kwak et al, 2015). Another alarming aspect of toxic behavior in these popular games is the risk that there may be a lot of younger kids playing them. It is likely that young children are playing popular games in which toxic behavior occurs, such as cyberbullying. Reports have shown cases of suicide in relation to cyberbullying (Pokin, 2015). League of Legends offers an alternative alias which results in anonymity for the player. This is often seen as being one of the reasons why players feel like they can exhibit toxic behavior since they do not feel like they can be held accountable for their toxic behavior (Kiesler, Siegel, and McGuire, 1984). Riot Game, on the other hand, argues that it is not the anonymity aspect that brings forth the toxicity but rather the lack of consequences for such behavior (Lin et al, 2013).

League of Legends is one of the most popular games in the world, and it is the top-grossing Free to Play MMO game on the market (Statista, 2017). LoL has a free to play business model with in-game microtransactions to support the game´s sustainability which is directly linked to the strength of the community. The stress caused by harassment or other types of toxic behavior can lead to players becoming fatigued and quitting the game. Leaving toxic behaviors unchecked can threaten entire communities to tear apart (Blackburn and Kwak, 2014). Toxicity is not something new, but with the growth of the gaming industry, the toxicity problem has grown with it (Lin et al, 2013). Tyack et al have researched retainment of players and found that out of around 50% of survey respondents (N = 760) that quit a MOBA game recently, 12,18% did so due to unsportsmanlike players. The participants of the survey were sourced from MOBA-related forums, online spaces, Facebook groups and Twitter etc (Tyack et al, 2016). Jeffrey Lin, neuroscientist and at the time game designer from Riot Games argued that social interaction in online multiplayer games determine if a player continues to play or quit. Lin states that players that are exposed to toxic behaviors are 320% more likely to quit the game entirely (Lin, 2015).

The earliest solutions to toxic behavior often required human intervention from a game master or admin. This can work for some games, but it is hard to scale it when the game has tens of millions of active players. This is why crowdsourced systems have been introduced, in which a lot of people can give their input. In 2011, Riot Games introduced The Tribunal, which was a system meant to deal with toxicity by using a crowdsourcing platform. The platform allowed rule-abiding players to analyze toxic players' behavior by looking at the chat logs and vote whether the player should be punished or not. This was a successful attempt, but it was far from a complete cure. Even though it did reduce the risk of recidivism, the tendency to relapse into a previous condition or mode of behavior, this kind of solution required a significant amount of human effort and the process is quite slow. There also has to be a number of reports sent in before the cases are even presented to the reviewers (Blackburn and Kwak, 2015). That is why this study aims to further research toxicity as an attempt to contribute towards the understanding of the subject area.

2 Introduction

In this paper, we are going to collect data from LoL games in order to examine the differences in toxic remarks between three major ranks. Data will be gathered from three different ranks, Bronze, Gold, and Diamond. The data collected from these games will be divided into five different categories, insults, taunts, complaints, spam and domain language.

2.1 MOBA

Multiplayer Online Battle Arena games referred to as MOBAs are played in independent games. Each game consists of ten players that play in teams of five. All game modes in MOBAs are competitive and highly reliant on team communication and collaboration. The primary goal for each team is to destroy the enemy team's base. It is achieved by destroying enemy players or enemy towers that protect the base. Players buy magical items to equip their character, they earn gold to buy items through destroying enemy players, neutral objectives, and enemy minions. Both teams have AI-controlled minions that earn players gold and are used to destroy enemy towers. Minions run in small groups that spawn from the team´s base. They run to kill the enemy team's base and try to do so until killed. Minions spawn throughout the entire game from set timestamps, usually every [30] seconds. Minions control the pace of the game, once they die near an enemy player, that player earns experience. A player will level up throughout the game to eventually reach the game´s maximum level. Every new level gives a player character upgraded base stats, such as more health.

Each team has a chat room where they have the option to communicate. Every player can choose to disable messages and in-game communication notifications from other players. Both teams share a chat room together where they can send messages to the enemy players. Players that play in the most common game modes usually communicate through sending messages, much like a Facebook chat room. Each player can send smart in-game notifications to other teammates to more efficiently coordinate their characters. Those types of notifications are often referred to as “pings”, they help players communicate without typing long messages.

MOBAs tend to have a few different game modes, the most commonly played game mode is often referred to as “ranked solo queue”. In the ranked solo queue players have the option to rank up through several different divisions that exist in the game. The division represents the player´s skill level, each player is matched with other players from the same division, to evenly balance the teams. Sometimes players are matched with players from higher or lower divisions. When that happens, players have the tendency to lash out and behave inappropriately. There are many factors that can lead to inappropriate behavior in MOBAs, it is, therefore, difficult to come up with solutions to solve issues regarding those behaviors.

3 Previous works within the subject area

Toxicity detection in multiplayer online games

Märtens et al. (2015) investigated the communication between players in highly competitive multiplayer online games. By using a language processing framework, they detected and analyzed profanity chat logs from a popular MOBA and develop a method to classify toxic remarks. Märtens et al. categorized all their collected data that they gathered from [12,000] Dota matches played in 2012. They focused primarily on chat logs rather than gameplay and defined toxicity as "the use of profane language by one player to insult or humiliate a different player in his own team" (p.3). They found that toxic remarks are more frequently preceded by kill events than random events. They also noticed that winning teams use less toxicity at the late stages of the match as it becomes apparent that they are victorious. On the contrary, the tendency to blame and use of toxic language was significantly higher for teams that were losing. Märtens et al, reach the conclusion that toxicity is fueled by the inherent competitiveness of MOBA games, but weakly linked to success. If players can be successful despite being toxic, they need a different incentive to stop insulting players and behaving properly.

Relating conversational topics and toxic behavior effects in a MOBA game

Neto, Alvino and Becker (2018) studied the effects of the conversational patterns used by players in LoL and relate them to different types of players (toxic, allies and enemies). They investigated whether those patterns affected player performance and player contamination. They found that the way that players tend to communicate with each other is directly linked to their performance and toxic contamination. Toxic player allies were more affected than their opponents, displaying lower performance and higher usage of toxic expressions. They observed that toxic player teammates became complicit in the toxic behavior, either by replicating or by tolerating it. Most of the time the toxicity that occurred was local, meaning that it tended to stay in the team rather than affecting the opponents. Opponents may have been affected by toxicity if the harassment was directly targeted at them. They also found that matches that are unbalanced tended to generate toxic behavior.

Exploring cyberbullying and other toxic behavior

Kwak et al. (2015) explored cyberbullying and other toxic behavior in team competition online games, specifically in LoL. They took advantage of the Tribunal system from LoL and analyzed ten million player reports on [1,46] million toxic players. Alongside the reports, they also saw the corresponding crowdsourced decision on whether or not the player in question should be punished or pardoned. They then compared the toxic behavior to theories from sociology and psychology to tried and explain why it occurred. They found that the bystander effect, being the tendency for observers to avoid helping a victim, particularly when they are immersed in a group, is evident when it comes to the reporting of toxic players. Few players after witnessing toxic behavior in a game actually report said player. The odds of them doing so are though greatly increased if it is requested for them to do so by another player. They also found connections between LoL and online communities such as Reddit, allowing anonymous user identities. Although the toxic behavior and cyberbullying between these two are different, the disconnect between the virtual world and reality is a common root that both cyberbullying and toxic playing stem from.

4 Definition of Toxicity

Riot defines toxicity as “any behavior that negatively impacts other players’ experiences” (Lin and Sun, 2013). Märtens et al. define toxicity as “the use of profane language by one player to insult or humiliate a different player in his own team”. As such, the use of “bad” words is a necessary, but not sufficient condition for toxicity (Märtens et al, 2015). Their study concludes that toxicity detection cannot be based on words alone but needs to take the context into account. Based on their research, we define toxicity as: Any behavior that negatively impacts other players’ experience. Furthermore, for this paper, the term toxic remark was constructed to assist while gathering data. Irony, sarcasm or humor cannot be objectively identified as toxicity due to the problematic aspect of interpretation. This paper will, therefore, disregard any form of toxic remarks related to irony, sarcasm or humor.

Toxic Remark: Any taunt, insult, complaint, spam or domain language directed towards another player or multiple players, regardless if that player/s interprets the remark as offensive or not.

Research Question:

How does toxicity change depending on rank in League of Legends?

5 Method

Toxic behavior has been defined, studied and researched in many different ways. The common understanding is that toxicity is a relatively broad term that can be interpreted to have multiple sub-categories, such as griefing, cyberbullying, mischief and cheating (Kwak et al, 2015). To properly research toxic behavior and how toxicity changes between different divisions in LoL, behaviors of toxicity was researched through in-game chat logs. A quantitative research approach has been used to examine toxicity through chat logs. All data was collected by using player observation. Fifteen games from rank Bronze, Gold, and Diamond were recorded during the data collection process for analysis. All results were then categorized and compared between the ranks. We can more accurately grasp how toxic each rank is by researching the amount of toxic remarks made in three different ranks. The three different ranks were chosen based off large skill gaps and knowledge. Players have different skill, domain language, and domain knowledge in these ranks, which creates different environments for toxicity.

We adopted a framework constructed by Neto, Alvino and Becker (2018) to identify toxic remarks using categorized clusters of toxicity in chat logs to count words. The counted words were compared between the divisions to identify the rank with the highest amount of toxic remarks. The adopted framework was customized specifically to research toxic remarks along with two new additions. Two categories were added to the framework to help the researchers identify the toxic remarks written in the chat logs. A category regarding spamming was added to ensure that cases where a single player spammed toxicity throughout the entire game did not contaminate the rest of the data. Therefore, whenever spamming occurred in a game, it was recorded but sorted into the spam category to avoid data diversions that would skew the final results. A category for domain language was also added, to ensure that specific cases where players used a lot of domain language to harass each other was recorded but recorded into its own category without contaminating the data that did not contain any domain language. The required domain language to understand such toxic remarks can be considered subjective due to differences in interpretation which is why it was made into its own category.

5.1 Toxic Remarks

The quantitative data was gathered by counting toxic remarks. This was accomplished by identifying toxic remarks. When identified, that specific remark was distributed into one out of five different categories. The categorization was based on the words used in the toxic remark. The categories represent clusters of toxic words put together in Neto, Alvino and Becker's framework (2018). They researched large bulks of in-game chat data to understand how players communicate with each other in relation to game performance and corroding team efforts. They looked at data collected from the tribunal and clustered words of similar meaning to create categories. These clusters were then analyzed and used to understand large bulks of in-game chat data efficiently. Their framework is composed of nine different categories relating to different words typed in the team and all-chat in LoL, seen in Figure 1.

Figure 1.

Table 1: Topics and Top-10 weighted words

Id.	Topic	Top-10 weighted words
0	*Other Langs.*	mira, vos, weon, ahora, wn, vida, qe, esa, eso, asi
1	*Tactics*	drag, ss, warded, red, flash, ward, eve, wards, tf, blue
2	*Tactics*	group, together, baron, fight, push, stay, focus, wait, lets, cait
3	*Tactic/Positive*	brb, mia, np, gj, thx, sry, ty, sorry, lag, mid
4	*Small-Talk*	op, dat, ap, blitz, ad, build, damage, lol, oh, xd
5	*Arguments*	talking, people, playing, mad, said, game, say, play, like, cause
6	*Complaints*	cs, fucking, adc, ur, support, stfu, lane, fed, idiot, dumb
7	*Complaints*	report, ban, troll, trolling, bg, reported, afk, feeding, pls, reporting
8	*Insults*	nooob, fu, noob, noobs, vs, ks, idiot, fck, stupid, omg
9	*Taunts*	ni***r, fa***t, c**t, mom, f*g, d**k, bi**ch, g*y, mad, a*s

Framework created by Neto, Alvino and Becker.

While Neto, Alvino and Becker were interested in discovering positive as well as negative communication it made sense to make categories for both outcomes. This study, on the other hand, uses only the last four categories from the framework, Complaints 6&7, Insults 8 and Taunts 9. The two additional categories were also used in combination, Spam and Domain language.

Descriptions for each of the categories used can be found in a list below. Insults, taunts and complaints were summarized based off Neto, Alvino and Becker's detailed explanations from their research.

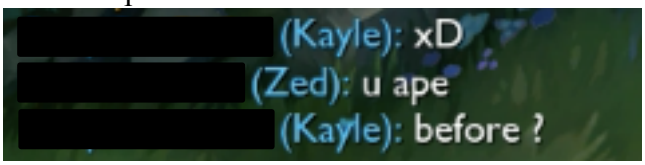
Insults: The insults refer to the alleged lack of skills of other players (e.g. noob and its variations), or possible mistakes during the game (e.g. idiot, stupid) (Neto, Alvino and Becker, 2018).

For example:



Taunts: Taunts are expressed mostly through racist (e.g. nigger), homophobic (e.g. faggot) and sexist (e.g. bitch) slurs. Neto, Alvino and Becker regard taunts as the most straightforward textual form of toxic behavior, as it tends to emotionally destabilize the other players. Four out of the five evaluators from Neto, Alvino and Becker's study agreed that the words in this category, compared to Insults, are significantly more offensive (e.g. hate speech, major insults).

For example:



Complaints: Complaints to other players are characterized by words expressing general dissatisfaction with other players and the match outcome. The majority of complaints have been shown to be about team issues. This might may be an example of non-intentional toxic behavior, as offenders may believe that their complaints are constructive. Some words indicate that players are not getting along, with the players blaming their teammates for possible mistakes that happened in the match. There are also players threatening to denounce others for toxic behavior. These complaints are not necessarily legitimate, as they can be just a part of the blaming cited before (Beckers et al, 2018).

For example:



Spam: Any insult, taunt, complaint or domain language used more than twice in a row by one or multiple players within a window of ten seconds. Meaning that if a player uses the same toxic remark more than twice in a row within a ten second time period, it will be classified as spam data.

For example:



Kayn repeatedly complained towards the bottom lane by spamming them with the “stop” word in combination with other toxic slurs. His complaints repeat to the point where it is classified as spam. Kayn continued to exhibit this behavior throughout the entire game.

Domain Language: Any insult, taunt, or complaint containing words that requires game knowledge to understand.

For example:



Interpretation difficulties may occur during the quantitative word counting approach. There will be moments where players typing toxic remarks appear as insults without any intention to

target a specific player. For example:

Player1: Shitt, fuck idiot
Player2: Who are you talking to? Raul or I?

Those are cases which will not be counted in this research as we strictly follow whether the remark is either; an insult, taunt, complaint, spam or domain language. Random insults without addressing a specific player will not be counted as a toxic remark, such as indirect harassments or other toxic slurs without context. Indirect harassment and remarks that are based on humor and sarcasm cannot be counted due to the complexity of subjective interpretation. Any case where the researchers cannot tell if it is an insult, taunt, complaint, spam or domain language directed at one or multiple players will not be counted as a toxic remark. All data comprised of toxic remarks were counted by the number of times a category occurred in each rank. For example; [100] taunts, [120] Complaints, [150] insults, [10] spam, [11] domain language (Rank: Gold).

All the counted toxic remarks was transcribed in word documents, all transcribed data is available in the bibliography. Any word or sentence that resembled positive communication was not transcribed in this research due to its lack of relevance. The toxic remarks counted were transcribed and timestamped. Any form of toxicity relating to the five different categories used in this paper were transcribed.

5.2 Participant Observation

To be able to gather both gameplay and chat logs from a LoL match, the researchers were required to be present in the game. Spectating a match that was previously played does not result in access to the chat logs. Dewalt (2001, p.100) argues that "For some, in fact, participant observation is an approach to getting deeper more solid contacts with people and situations rather than a method itself". Participant observation will be applied by playing LoL matches while recording the gameplay. In addition to recording, the researchers will be jotting down timestamps throughout the game. The timestamps will be indications of when a toxic remark is stated in the game. This remark is one of five indications, that an insult, taunt, complaint, spam or domain language have been stated by a player directed towards another player or multiple players. The notes will improve the efficiency of analyzing the games in the later stages. Instead of having to look at an entire LoL game in real-time it is now possible to fast forward with the guidance of the timestamps taken.

5.3 Code of Conduct

In order to apply participant observation efficiently, a code of conduct was outlined for the researchers while participating in the games. The three different ranks, Bronze, Gold, and Diamond were divided between the two researchers based on their previous experience. One researcher (A) had six years of experience playing LoL and is currently in the Diamond rank, the other researcher (B) had four experience of playing LoL is currently in the Gold and Platinum ranks. The researchers played each of the [15] games corresponding to their previous experience. Researcher B played all games in Bronze and Gold, and researcher A played all games in Diamond. All games were played with set limitations. Those limitations included:

Bronze

Played only the support role and performed all actions that were expected from the role but without using advanced techniques that could potentially carry the game or skew the outcome of the game to a large degree. All games were played on accounts corresponding to the associated rank; in this case the account was visibly seen as Bronze Rank III to other players.

Gold
Played any role and to the best of his ability. Played any champion that the researcher had prior experience with. All games were played on accounts corresponding to the associated rank; in this case the account was visibly seen as Gold Rank IV to other players.

Diamond
Played any role to the best of his ability. Played any champion that the researcher had prior experience with. All games were played on accounts corresponding to the associated rank; in this case the account was visibly seen as Diamond Rank III to other players.

1. The researcher never took the initiative to start a conversation with other players in the game. On the other hand, if a player in the game reached out to the researcher they were allowed to reply. It was allowed to excuse small misplays and mistakes instead of staying quiet.

2. The researcher was not allowed to reply with the intent of bringing forth toxicity. This includes being truthful and not using profanity.

3. It was also not allowed to play in a certain way with the intention to trigger other teammates.

4. The researcher never gave up in a game and never stopped trying to win. Riot Games' Summoner's Code was adopted and followed accordingly (Riot Games, 2017, Summoner's Code).

5.4 Limitations

1. Sarcasm and humor used as harassment in the chat log towards other players will not be counted as a toxic remark.

2. Insults, taunts, or complaints where the target behind the remark was unclear to the researchers will be disqualified as a toxic remark.

6 Results

All results have been summarized into two different tables and one column chart. The column chart visualizes how many toxic remarks each rank had complemented with the ranks respective standard deviation. The first table represents all data collected and compiled from the three ranks. The second table represents the mean average and standard deviation for toxic remarks and all the individual categories All data is available in the bibliography for a more advanced inspection of the data collected. The recorded games have been destroyed to protect anonymity.

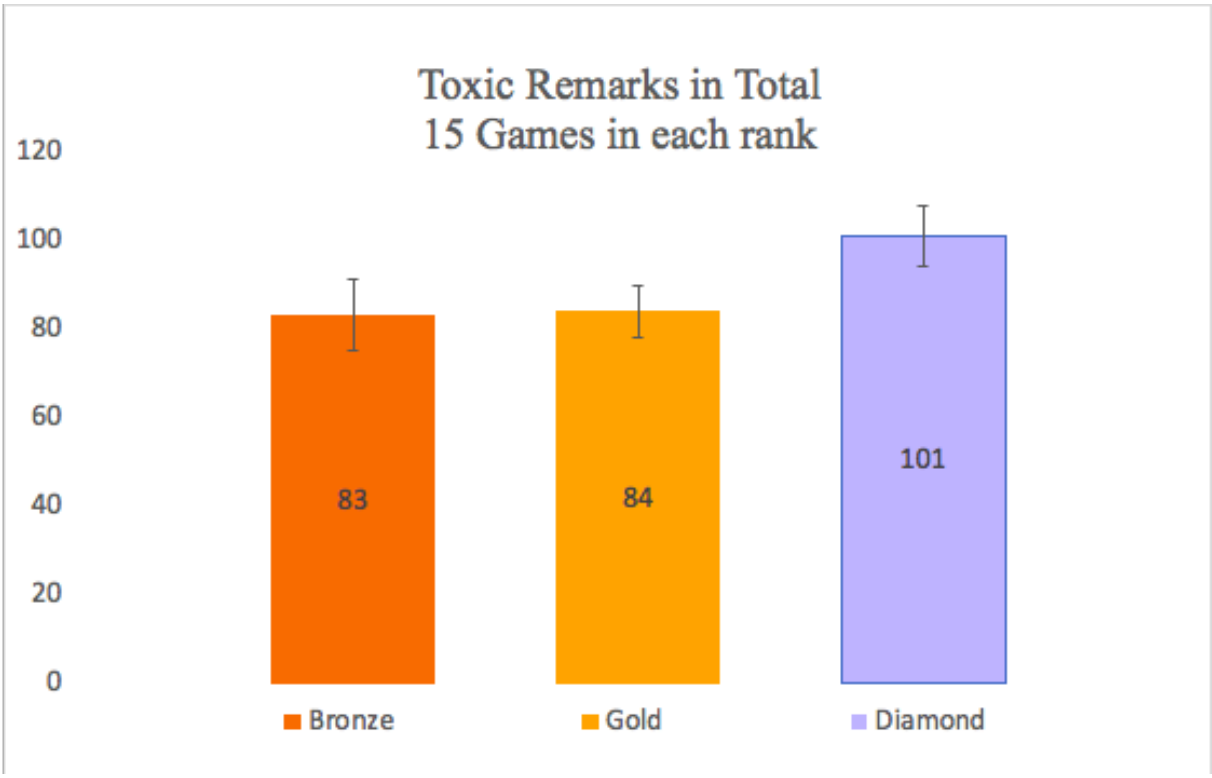


Table 1. Summary of Results

Total Stats	Games played	Wins	Losses	Win Rate (%)	Toxic Remarks	Games free of toxic remarks	Insults	Complaints	Taunts	Domain Language	Spam
Total	45	21	24	47	268	9	56	102	23	85	0
Bronze	15	8	7	53	83	3	22	42	9	10	0
Gold	15	6	9	40	84	2	18	33	4	28	0
Diamond	15	7	8	47	101	4	16	27	10	47	1

Total results of player toxicity within each rank

Table 2. Mean Averages of Results

Mean Average (SD)	Toxic remarks	Insults	Complaints	Domain Language
Total	6.0 (7.3)	1.2 (1.9)	2.3 (3.6)	1.9 (4.6)
Bronze	5.5 (8.6)	1.5 (2.5)	2.8 (4.9)	0.7 (1.1)
Gold	5.6 (6.0)	1.2 (1.8)	2.2 (2.9)	1.9 (2.2)
Diamond	6.7 (7.1)	1.1 (1.7)	1.8 (2.1)	3.1 (3.3)

Mean (standard deviation) of instances of toxicity within each rank

7 Analysis

Riot Games argues that it is the lack of consequences that brings forth toxicity rather than players' freedom of anonymity (Lin and Sun, 2013). Consequences or not, toxicity still has its place in the modern sports era of 2019. Look at boxing, the largest showcase matches between two famous boxers rely on toxicity to promote and advertise their big showmatch. The boxers are pitted against each other on stage during press conferences where both boxers showcase their strength, weight and previous titles. During such conferences, insults and taunts are being thrown by the boxers at each other to attract more fans to the big match by creating drama. Competitiveness in sports will most likely always have a toxic side to it, whether athletes or gamers want it or not. This paper investigated toxicity by examining three different ranks in LoL. The purpose was to understand why players from the different ranks use the language they use in relation to their rank. Why is Bronze players using a higher frequency of complaints for example, compared to Diamond and Gold players? This leads into our research question: How does toxicity change depending on rank in LoL?

The results from the data collection showed that the Diamond rank contained more toxic remarks than Gold rank and Bronze rank, with a total of [101] toxic remarks over [15] games. Gold and Bronze showed similar results in terms of overall toxic remarks counted when compared. It was counted [83] remarks for Bronze and [84] remarks for Gold. The results showed that lower ranked players communicate their toxicity differently than higher ranked players. Bronze players tend to use complaints while directing their concerns to their teammates, [42] counted. Gold players use a wider spread of toxicity when using remarks to harass other players, having moderate to high counts in each of the categories. Diamond players rely on advanced domain knowledge to communicate, as a result, [47] out of [101] (46.5%) toxic remarks were counted as domain language.

The result showed that the mean average for a toxic remark to appear in a ranked game in Bronze, Gold, and Diamond was six with a standard deviation of [7.1]. By playing in those ranks players can expect to encounter six toxic remarks directed at one or multiple players every game. That means that every game is expected to contain some sort of toxic behavior, when averaging a mean for toxic remarks in LoL. Data shows that four out of [15] games in Diamond were free of toxic remarks, two out of [15] in Gold and three out [15] in Bronze. This indicates that every Diamond game has higher toxic remark ratio once players start calling out each others mistakes. This statement is true because the diamond rank still

contained the highest amount of toxic remarks [101] while having the lowest amount of toxic games [11]. The Diamond games risk being harsher while having a higher ratio of toxic free games. This statement could be disregarded due to the low difference in amount of toxic free games, Diamond [4] games compared to Bronze [3] games, further research might be needed to prove this statement.

Results showed that nine games out of [45] were free from any toxic remarks. Unfortunately, we cannot claim that they were completely free from toxic remarks. From our perspective, we registered no toxic remarks either in team chat nor all chat, but we cannot know what was written in the enemy team's chat logs. This goes back to the theory Märtens et al. brings up in their study 2015, that losing teams tend to be more toxic than winning teams in the later stages of the game. Based on Märtens findings, those games would most likely have contained toxicity, at least the chat log from the losing team. However, four out of nine games never reached the later stages of the game. We speculate that the games ended so quickly due to one team being too strong and therefore it reduced the amount of toxic remarks since the losing team never saw the possibility of winning. As a competitive player, it would not make sense to try and motivate your teammates if no one in the team felt like there is a chance of winning. Therefore, many players instead prefer to lose the game as fast as possible to move on to another game in which they feel like they can win. The word "ff 15" is common toxic domain language which can be noted in the corpus. The remark simply means that that player intends to lose the game as fast as possible, which is demoralizing for teammates in the game, still trying to win.

Only one of the toxic free games resulted in a defeat, compared to the other eight victories that were toxic free. That means that we only had one loss that was toxic free out of [45] total games played. Out of these nine toxic free games, five of them were highly one-sided, meaning one team outperformed the other. Considering most of the players in the opposite team got outperformed it may have influenced the amount of toxic remarks, no one felt like they had the "right" to be toxic towards one another. Generally, from our experience, when a player is being toxic in LoL they are usually performing above expectations. Player performance often increases the player's self-confidence, it makes them arrogant and highly ignorant to rational arguments. This increases the risk of influencing teammates in a negative manner if ever questioned or criticized by another teammate. In our case it seems like the prior theory is more accurate considering we did not register any toxic remarks during these nine games. On the other hand, there were games where players on the opposite team did perform well. In these cases, it is safe to assume that these players either chose to express their toxic remarks in the team chat log or that these players were not toxic players and took the loss like an honorable loser.

The results from the data collection showed that spam only occurred once in of all the games played. The outlying result revealed a concerning discrepancy in the data, and without the category to separate the data into, everything could have fallen apart. This was replicated based on Märtens et al. (2015) previous research. They used a spam category to filter out diverging data that could risk skewing the results from the other categories.

Results showed that Bronze players used [10] toxic remarks as domain language, Gold players used [28], and Diamond players used [48]. That means that higher ranked players tend to use more domain language. Domain language is connected to player skill, and player skill to player rank. A higher ranked player would supposedly have more knowledge about the

game compared to a lower ranked player, which is why they would tend to use remarks linked to domain language. This is an interesting aspect identified from the results, because this can be linked to the perception of griefing, referring to whether a player feels griefed by another player or not. If a low ranked player played together with a high ranked player, a domain knowledge gap would exist between the players. The higher ranked player would perceive some plays to be griefing by the lower ranked player. The lower ranked player would not perceive the same plays as griefing due to his lack of domain knowledge. Despite arguments regarding intention or not, the higher ranked player would most likely perceive the lower ranked player to be griefing him/her if they played together in a competitive environment.

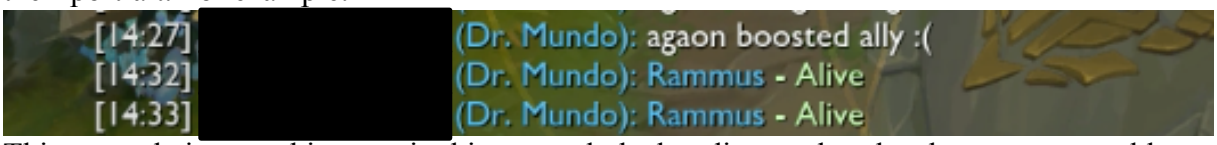
Discussion

The Ping System

Throughout the data collection process, we noticed that player behaviors have evolved with new systems when interacting with each other. The pinging system was introduced to make quick communication even quicker. It allows players to, with one click and drag, communicate to their team to fall back, if they are on their way to the fight etc. For this purpose, the system works very well. Unfortunately, there are also downsides with the pinging system. Parts of the pinging system is used to trigger other players by misusing the system in a toxic manner. The pinging system also allows players to ping other player's portraits and display in the chat if "The pinged player is alive/dead". This is often used extensively when someone makes small mistakes and other players want to highlight this mistake. We believe that most LoL players are aware of this issue and it functions like a call out for mistakes. In many of the games played, this action was used to call out people's mistakes which often leads into hectic discussions between players.

Indicators

To improve the methodology of this study we believe that the pinging system can be integrated with the way we worked to increase the total amount of data that can be collected. We worked immensely with relating toxic remarks to a certain player. Instead of only looking at the written chat logs, we believe that by looking at pings as indicators it can result in more data. It was very common when playing the games that players would write toxic remarks but instead of writing the name of the target the remark was intended for, they instead pinged their portrait. For example:



This example is something we, in this research, had to disregard as data because we could not determine to whom Dr. Mundo were talking to. If a future study with similar aims as ours would be made we would suggest including the pinging system as a way of identifying targets of toxic remarks. Pings are frequently used by players playing LoL at the moment and would add on a lot to the data collection process.

Proxy Singed

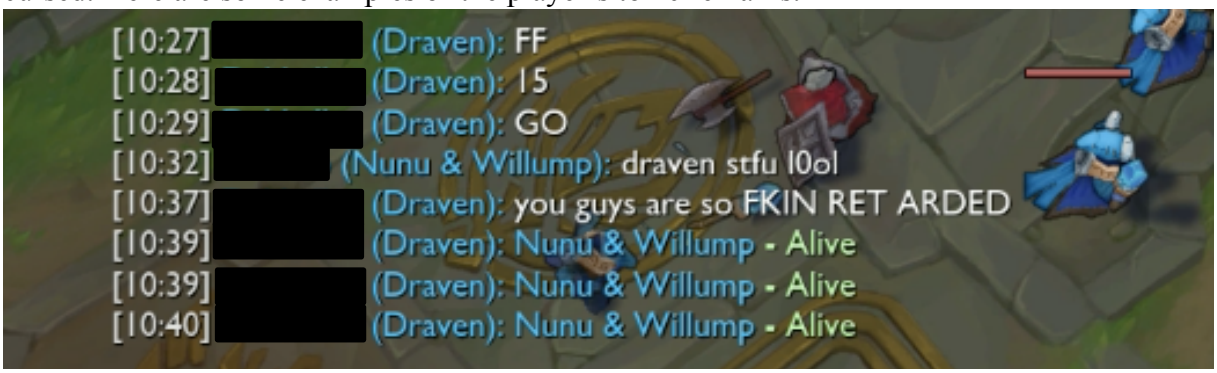
Domain knowledge is an important aspect of this paper, not only when it comes to toxic remarks and how to indicate if they are meant to be toxic, but also to determine what the community judge as accepted behavior or playstyle. So, what if a playstyle revolved around

dying continuously? There have been many strange playstyles in LoL since we started playing, but the one that stands out is Proxy Singed.

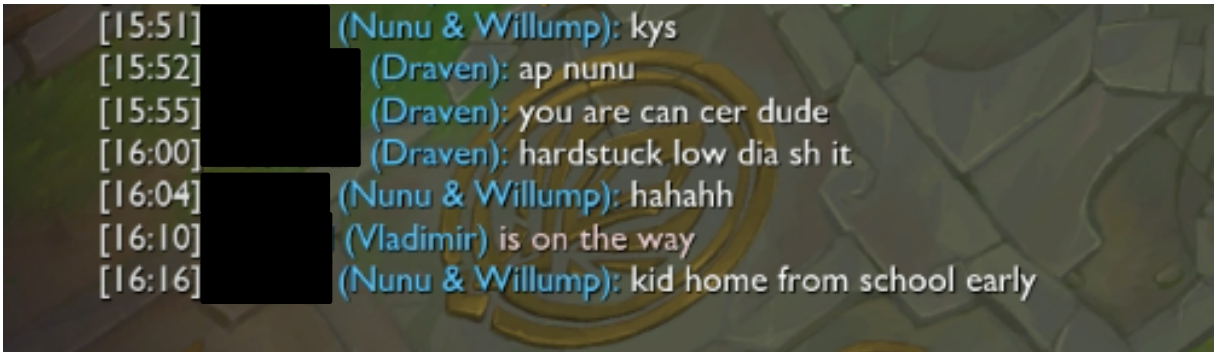
Proxy Singed is something only a well-versed Singed player can use effectively and to his team's benefit. The playstyle is played by positioning Singed behind enemy towers and killing their minions "behind enemy lines" more or less. The goal, if a player is playing Proxy Singed, is to attract the enemy team's jungler's presence and by doing so relieving their team of potentially devastating ganks. This playstyle often results in a lot of deaths for the Singed player, but it works decently with the help of one mechanic existing in LoL. The more deaths a player has without ever getting a kill on an enemy player, the less gold they are worth when they die. For example, killing a Singed with six or more deaths gives very little reward, especially since killing a Singed is never easy. There is a reason for why there is a saying in LoL that goes "Never chase a Singed" (Esportspedia, 2019). We always imagined how people reacted when they first saw a Proxy Singed back in 2013, 2014 when the playstyle was not common domain knowledge yet. Interestingly the playstyle is so in-depth that it requires a certain amount of domain knowledge to understand why it is even good in the first place. For this playstyle to become an "acceptable playstyle" to the community of LoL it would have to become a meta (most effective tactics available) within LoL. If the playstyle is not considered meta, any player without the domain knowledge to understand why it is good would report people playing Proxy Singed. These players would be reviewed for punishment because other player's domain knowledge is too small.

The Draven Case

We found one specific case where a player exhibited an excessively toxic personality. The specific way the player communicated suggested that the player know how to behave in a toxic manner without being punished. The player displayed a toxic personality after getting killed two versus one around minute [10.00], and after that the game, in his eyes, were over. However, it was not the fact that the player was cursing in chat but rather how the player cursed. Here are some examples of the player's toxic remarks.







If we look at how the player is constructing their sentences it is evident that there is some kind of pattern to it. The player is deliberately making spacings mid-word when there is a toxic word involved. We found this very interesting and this was mainly because neither of us had ever come across such behavior before. Discussing it between ourselves our speculations was that the player either knew or thought they knew how to avoid being banned for using toxic words. We would be immensely surprised if Riot Games did not have a way of detecting this simple but innovative and easy way of ducking the banning system.

Alternative methods

Participant observation, to us, turned out to be the optimal method when conducting this study. Considering the given time scope, we figured that even though we would have to play through plenty of LoL games it would not be a problem. However, the use of participant observation had to be slightly altered. Initially it was meant that one of us would be the participant in game while the other observed the gameplay. By doing it this way the researcher that was participating in the game could 100% focus on just playing the game. The other researcher would observe everything that could be relevant to our paper and jot down each timestamp. We truly believe that this would be the superior way of using Participant observation, but due to time restraints, we had to play at the same time in order to gather all the games, which is why we changed the method. All of the games were going to be recorded anyways, which meant that analysis and observation of toxicity could always be performed by both researchers after the game. Playing games and collecting data was of higher priority rather than being 100% in the game as a participant.

The study could be argued to be invalid ecologically when considering the method of participant observation. However, alternatives such as spectating LoL games through twitch as a possible method was discussed prior. Initially, the idea was to spectate games in LoL's client, but it proved to be problematic since the replays of games through LoL's client does not display the team's chat log. We then discussed whether or not observing games via Twitch or streamers could work. This did seem more promising due to the opportunity to observe the chat logs, but we also saw small problematic aspects that could come to ruin the research. Relying on data collection from a Twitch streamer's LoL games could prove to be of high risk. First and foremost, we would have to find three different streamers that belong to each of the ranks that we wanted to investigate, Bronze, Gold, and Diamond. Then, there is the muting aspect. It is common for players, especially players that want to climb higher in ranking, to mute players that are too toxic too handle. Instead of listening to the toxic player for the entirety of the game, streamers tend to mute these players. It is also highly common for streamers to be identified by fans who play the game, those fans tend to seek for the streamer's attention, influencing the chat log data. So not only are we dealing with the

streamer muting people that could add on to the data collection, but also the streamer. Being in the game, could potentially risk influence the behaviors of the other players in the game.

The code of conduct and limitations prevented us from influencing the game to a large degree compared the alternative of using a streamer. We would in other words have to find a streamer that we know for a fact is not a toxic player, otherwise every game they played would exhibit incredibly toxic behaviors that would skew the reality of the situation. The only realistic alternative to participant observation would be through observing a third-party participant from corresponding rank being researched. That would imply finding suitable players for each rank, without any of them having higher experience of ranked play than the rank in research. It is more than difficult to find players playing in Bronze that are available for such studies as this, than to find oceans on land.

Another aspect of the participant observation that could be disputed is if whether our performance influenced our teammate's performance and behavior or not. Based on our code of conduct and limitations considering the method used, yes, it is unavoidable to not influence a game in LoL as a player. However, the measures to reduce influence have been taken to an extent where teammates only could be influenced through horrific performance from the player researching. One limitation constricted us from playing champions that we had not experience with prior, meaning that we only played champions we were good at, essentially reducing the risk of us playing horrifically.

Conclusion

The research found that the Diamond rank has a higher frequency of toxic remarks (101) compared to the other two researched ranks, Bronze (83), and Gold (84). It was found that there is direct link between player expertise and the use of domain language when using toxic remarks directed at other players. Bronze players tend to use complaints and insults when remarking their teammates while Gold players tend to use a wider spread of toxic remarks. Diamond players tend to use more domain language when remarking their teammates. Domain language is used by players that have experience from playing the game, it is directly tied to game knowledge. This means that higher ranked players rely on toxicity that requires game knowledge to understand, while lower ranked players rather use more simple complaints and insults. That may be because they know that anyone can understand such remarks without having prior knowledge about the game. Getting the message across to the teammates might be of higher value compared to simply harassing without purpose, for lower ranked players.

Our aim with this paper was to further investigate toxicity in LoL more deeply and try to find the reasons behind why players communicate the way they do. As a more personal aim to the researchers we were curious as to which rank would contain the highest amount of toxic remarks and understand why that is. Conclusions such as the fact that higher ranked players tend to use more domain language when being toxic is a very interesting finding. We are certain there are more findings to be discovered within LoL's toxicity problem.

References

1. Kwak, Haewoon, Jeremy Blackburn, and Seungyeop Han. “Exploring Cyberbullying and Other Toxic Behavior in Team Competition Online Games.” ArXiv:1504.02305 [Physics], April 9, 2015. http://arxiv.org/abs/1504.02305.

2. Pokin, S, 2015, Pokin Around: The Story of Megan Meier's suicide, Spring Field News-Leader [Online], 27th of May, Available at: https://eu.news-leader.com/story/life/2014/11/19/pokin-around-biggest-story-young-girls-suicide/19291825/ [Accessed 17th April 2019]

3. Kiesler, S., Siegel, J., & McGuire, T. W. (1984). Social psychological aspects of computer-mediated communication. American Psychologist, 39(10), 1123-1134. http://dx.doi.org/ 10.1037/0003-066X.39.10.1123

4. Riot Games, 2013. *Improving Player Behavior in League of Legends*. [press release] 30 March 2013. Available at: https://www.youtube.com/watch?v=a-5yKdpR0kU&t=71s [24th April].

5. Statista (2019). Topic: League of Legends. [online] Available at: https://www.statista.com/topics/4266/league-of-legends/ [Accessed 23 May 2019].

6. Blackburn, Jeremy, and Haewoon Kwak. “STFU NOOB! Predicting Crowdsourced Decisions on Toxic Behavior in Online Games.” *ArXiv:1404.5905 [Physics]*, April 23, 2014. http://arxiv.org/abs/1404.5905.

7. Tyack, April, Peta Wyeth, and Daniel Johnson. “The Appeal of MOBA Games: What Makes People Start, Stay, and Stop.” In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play - CHI PLAY ’16*, 313–25. Austin, Texas, USA: ACM Press, 2016. https://doi.org/10.1145/2967934.2968098.

8. Riot Games, 2015. *More Science Behind Shaping Player Behavior in Online Games*. [press release] 2-6 March 2015. Available at: https://www.gdcvault.com/play/1022160/More-Science-Behind-Shaping-Player [Accessed 24th April].

9. Martens, Marcus, Siqi Shen, Alexandru Iosup, and Fernando Kuipers. “Toxicity Detection in Multiplayer Online Games.” In *2015 International Workshop on Network and Systems Support for Games (NetGames)*, 1–6. Zagreb, Croatia: IEEE, 2015. https://doi.org/10.1109/NetGames.2015.7382991.

10. Mesquita Neto, Joaquim Alvino de, and Karin Becker. “Relating Conversational Topics and Toxic Behavior Effects in a MOBA Game.” *Entertainment Computing* 26 (May 1, 2018): 10–29. https://doi.org/10.1016/j.entcom.2017.12.004.

11. Lin, Holin, and Chuen-Tsai Sun. “The ‘White-Eyed’ Player Culture: Grief Play and Construction of Deviance in MMORPGs,” n.d., 12.

12. “Participant Observation: A Guide for Fieldworkers - Kathleen Musante (DeWalt), Billie R. DeWalt - Google Books.” [Accessed 29th April 2019], 2019. https://books.google.se/books?hl=en&lr=&id=ymJJUkR7s3UC&oi=fnd&pg=PP1&dq=dewalt+dewalt&ots=Uhsq78BqS9&sig=jBjihomBG68AFYdjO2owS_bebmM&redir_esc=y#v=onepage&q=dewalt%20dewalt&f=false.

13.Riot Games, 2017. The Summoner’s Code. [Website] Available at: https://na.leagueoflegends.com/en/featured/summoners-code [Accessed May 1th].

14. EsportsPedia, 2019. Singed/History. [Website] (May 13th 2019) Available at: https://www.esportspedia.com/lol/Singed/History [Accessed May 15th].

Appendix

A. Bronze
https://docs.google.com/document/d/1Xu3zNnNub1OXAK5gS2X0vnQo6DOIHah7M3TnWkjMMFk/edit?usp=sharing

B. Diamond
https://docs.google.com/document/d/19P04hYZEQwuQWev3Z_JW6VpCLFOT25KNFsQvgWTURvI/edit?usp=sharing

C. Gold
https://docs.google.com/document/d/18qrgwnaxOKaeg_SNEKi0hpPv9NSaeQntL7F9OwbeZ9k/edit?usp=sharing