석 사 학 위 논 문
Master's Thesis

예측 위반 이론의 적용을 통한 온라인 멀티플레이어 게임에서의 부정적인 팀원 인식 분석

Application of Expectancy Violation Theory in Understanding Negative Teammate Perception in Online Multiplayer Gaming

2024

이 주 훈 (李 珠 薰 Lee, Juhoon)

한 국 과 학 기 술 원

Korea Advanced Institute of Science and Technology

석 사 학 위 논 문

예측 위반 이론의 적용을 통한 온라인 멀티플레이어 게임에서의 부정적인 팀원 인식 분석

2024

이 주 훈

한 국 과 학 기 술 원

전산학부

예측 위반 이론의 적용을 통한 온라인 멀티플레이어 게임에서의 부정적인 팀원 인식 분석

이 주 훈

위 논문은 한국과학기술원 석사학위논문으로
학위논문 심사위원회의 심사를 통과하였음

2023년 12월 27일

심사위원장 장 정 우 (인)

심 사 위 원 김 주 호 (인)

심 사 위 원 도 영 임 (인)

Application of Expectancy Violation Theory in Understanding Negative Teammate Perception in Online Multiplayer Gaming

Juhoon Lee

Major Advisor: Jeong-woo Jang
Co-Advisor: Juho Kim

A dissertation submitted to the faculty of
Korea Advanced Institute of Science and Technology in
partial fulfillment of the requirements for the degree of
Master of Science in Computer Science

Daejeon, Korea
December 27, 2023

Approved by

Jeong-woo Jang
Professor of Digital Humanities and Computational Social Sciences

The study was conducted in accordance with Code of Research Ethics[1].

[1] Declaration of Ethical Conduct in Research: I, as a graduate student of Korea Advanced Institute of Science and Technology, hereby declare that I have not committed any act that may damage the credibility of my research. This includes, but is not limited to, falsification, thesis written by someone else, distortion of research findings, and plagiarism. I confirm that my thesis contains honest conclusions based on my own careful research under the guidance of my advisor.

MCS

이주훈. 예측 위반 이론의 적용을 통한 온라인 멀티플레이어 게임에서의 부정적인 팀원 인식 분석. 전산학부 . 2024년. 34+iv 쪽. 지도교수: 장정우, 김주호. (영문 논문)
Juhoon Lee. Application of Expectancy Violation Theory in Understanding Negative Teammate Perception in Online Multiplayer Gaming. School of Computing . 2024. 34+iv pages. Advisor: Jeong-woo Jang, Juho Kim. (Text in English)

초 록

온라인 팀 기반 멀티플레이어 게임은 플레이어 간의 갈등으로 인해 플레이어와 게임 커뮤니티에 큰 피해를 야기한다. 이 연구는 예측 위반 이론을 적용하여 플레이어들이 팀원의 성과에 대한 기대의 부정적 위반이 팀 내 갈등에 어떻게 기여하는지를 탐구한다. 본 연구에서는 플레이어가 게임 내 주어지는 랭크라는 잣대를 기반으로 팀원의 기대치를 조율하는 데 어려움을 겪는다는 것을 보인다. 또한 기대가 위반되는 경우 플레이어들은 팀원에 대한 부정적인 감정적 반응을 경험한다는 것을 보인다. 이러한 결과를 통해 팀 기반 온라인 게임에서 플레이어의 적절한 기대치 설정에 대한 중요성과 동시에 이를 이루는데의 어려움을 다루며, 플레이어 기대 수준의 개선을 통해 게임 경험을 향상시키기 위한 시사점을 제안한다.

핵 심 낱 말 예측 위반 이론, 기대 위반 이론, 온라인 게임, 갈등, 팀내 갈등, 플레이어 인식

Abstract

Team-based online multiplayer games are faced with a pervasive challenge of within-team conflicts. We apply the Expectancy Violation Theory (EVT) framework to explore how negative violations of players' expectations towards their teammates' performance contribute to within-team conflict. The study employs a between-subjects survey methodology to measure expectancy and its subsequent impact on emotional and attributional perceptions. Our research reveals that players struggle to align their pre-interaction expectations set by rank. Additionally, regardless of how expectations are violated, players experience heightened negative emotional responses towards their teammates. The findings contribute to the understanding of player expectations in team-based gaming contexts, highlighting the challenges in proper expectation setting. The study suggests implications for game designers and developers to address these issues and enhance the overall gaming experience in team-based online multiplayer games.

Keywords Expectancy Violation Theory, Online Gaming, Team Conflict, Player Perception

Contents

Contents i
List of Tables iii
List of Figures iv

Chapter 1. Introduction **1**

Chapter 2. Expectancy Violation Theory **3**
2.1 Key Concepts 3
2.1.1 Expectancy 3
2.1.2 Violation and Valence 4
2.1.3 Violation Response 4
2.2 Application of Expectancy Violation Theory to *League of Legends* 5
2.2.1 Research Context: League of Legends 5
2.2.2 Expectancy Violation Theory in Group and Collaborative Contexts 5
2.2.3 Expectations and Violations in *League of Legends* . . . 6
2.3 Role of Rank in Player Expectations 7

Chapter 3. Hypotheses **8**
3.1 H1: The Misalignment of Expectations on Rank 8
3.2 H2: The Effect of Expectations of Rank on Negative Expectancy 8
3.3 H3: The Effect of Negative Violation on Player Perception . . . 9

Chapter 4. Methodology **10**
4.1 Data Collection 10
4.1.1 Participants 10
4.1.2 Game Selection 10
4.2 Procedure 11
4.2.1 Participants 11
4.2.2 Task 1: Investigating Expectation Alignment of Ranks . 11
4.2.3 Task 2: Investigating Effect of Negative Expectancy on and Player Perception 13

Chapter 5. Results **16**
5.1 Rank Prediction 16
5.1.1 Perfomance Prediction for Teammates 17
5.1.2 Qualitative Analysis 17
5.2 Effect of Rank on Expectancy 20
5.2.1 Qualitative Analysis 20
5.3 Negative Violation and Player Perception 22

Chapter 6. Discussion **25**
6.1 Lack of Consistent Expectations for Rank 25
6.2 Unrealiblilty of Rank as an Expectation Setter 25
6.3 Impact of Negative Violations in Online Gaming 25
6.4 Implications for Platforms and Players 26

Chapter 7. Limitations and Future Work **27**

Chapter 8. Conclusion **28**

Bibliography **29**

Acknowledgments **33**

Curriculum Vitae **34**

List of Tables

4.1 Demographic information of the study participants 12
4.2 Videos selection for Task 1 15
4.3 Video selection for Task 2 15

5.1 Interquartile range (IQR) and median of predictions for each rank 17
5.2 Participation distribution in Task 2 for each condition by expectation direction from Task 1 21
5.3 Kruskal-Wallis test results for expectancy for each expectation direction group 21
5.4 Count of sample points of each expectancy 23
5.5 Kruskal-Wallis test results for expectancy and emotion 24
5.6 Posthoc Dunn Test Results for Negative Violation and Confirmation for emotion 24

List of Figures

5.1 Distribution of rank predictions by rank 19
5.2 Distribution of rank predictions by rank and win condition 19
5.3 Confidence level of rank prediction by rank 19
5.4 Quadrant plot of the expectancy distribution 23

Chapter 1. Introduction

Team-based online multiplayer games are ripe with within-team conflicts. Popular games spanning various genres and themes, such as *League of Legends* [31], *Dota 2* [51], *Overwatch* [29], and *Counter-Strike: Global Offensive* [27], all struggle with moderating within-team conflict amongst their multi-million player base [6]. In extreme cases, such conflicts between the team members can lead to toxic behaviors, such as flaming, sabotaging, and griefing [36, 40, 30, 26], undermining both the overall gaming experience and the well-being of the individuals involved [43, 46, 37]. Such prevalence of hostility in the gaming environment cuts into the profit of the game companies, who must deal with player exodus [13] and invest heavily in player moderation [6]. Thus, game designers must gain a better understanding of why such within-team conflicts occur to mitigate their rampancy and effect.

Previous research has revealed that these conflicts and toxic behavior are prevalent in certain in-game situations. While some research has focused on persistent reasons, such as how personality (such as extraversion [35] and impulsivity [41]) or normalization of toxic behavior [9] correlates with an increased chance of hostility, other studies and statistics show that the more immediate cause does not lie with permanent, static individual characteristics but rather by their response to certain stimuli — players are often triggered into visceral reactions when experiencing a cycle of emotional turmoil and negative conduct [38] or when facing specific events such as being killed [42]. It is especially noteworthy that in *League of Legends*, a widely played Multiplayer Online Battle Arena (MOBA) game, 95% of players exhibit toxicity intermittently, typically when provoked by in-game events. Remarkably, these occasional toxic players account for 86% of in-game reports [14], showing that players are not the cause of conflict, but rather reactants to negatively affecting stimulus. This also aligns with research looking at the causes of other deviant behavior (e.g., trolling [24]) online. We need to understand how such stimuli may be causing players to experience negative emotions and exhibit negative behavior.

This thesis posits that such negative responses are correlated with negative violations of a player's pre-interaction expectations towards another teammate's performance, based on the Expectancy Violation Theory (EVT) framework by Burgoon (1978) [16]. We approach these situations, such as kill events, as potential negative violations of pre-formed expectations of how a teammate should perform or behave. Works in the domain of EVT have shown that violation of an interactant's expectations in an individual-to-individual interaction leads to stronger arousal and social judgment compared to confirmation. For this study, we explore this idea in the context of *League of Legends*, a team-based, competitive battle arena game whose defining characteristics of high interdependence and temporary team formation are grounds for players to form pre-game expectations of their teammates' performance. Concerningly, these characteristics are common in other battle arena games and other competitive games beyond *League of Legends*.

We specifically narrow the scope of the expectation to individual teammate-to-teammate basis based on *rank*, a standard performance metric provided by the game. The rank, as an all-encompassing and most visible label, may serve as a pre-interactional communicator characteristic that critically shapes the player's perceptions toward another individual's performance in the group. While the previous works on EVT have mostly focused on communication and partnership-based offline human-to-human interactions [45, 25]. However, more recent explorations have extended the framework into computer-mediated communication (CMC), especially in understanding individual expectations toward another group member [39, 47] in collaborative [34] contexts. We note the similarity of such domains and *League of Legends* and denote that despite its contexts, we focus on the dyadic interaction between two teammates in the game.

Through such lens of EVT in the team-based gaming context, we aim to answer how a player's expectations towards their teammate are formed, how such expectations may be negatively violated, and how these negative violations can lead to more negative perceptions and responses towards the teammate. To do so, we conduct a between-subjects survey study measuring how accurately a player forms expectations given rank, whether setting higher expectations through rank will result in higher negative violations, and finally, whether such negative violations can hurt the emotional and attributional perception of the violator. We also discuss through what perspectives and factors a player views their teammates' performance through a qualitative analysis of the survey responses. The results show that players are unable to set aligned expectations on rank, contributing to a lack of violations caused by rank differences. However, regardless of how the expectations have been set, the violation of such expectations intensified the negative emotional response of the players toward their teammates. We further discuss how these results may suggest difficulties in proper expectation setting in online multiplayer gaming contexts.

The thesis makes the following major contributions:

- A demonstration of misalignment of players' expectations of a teammate based on rank
- An understanding of how players form expectations of a teammate's performance
- An exploration of how negative violation leads to unfavorable perceptions, contributing to a more negative and hostile environment

Chapter 2. Expectancy Violation Theory

Expectancy Violation Theory (EVT) was first proposed by Burgoon (1978) to explain and categorize the consequences of unexpected behavior in interpersonal communication [16]. While the theory originally explored nonverbal proxemic distance violations, or how close a person is standing to the individual, it has been since extended to a myriad of other contexts and behaviors, including emotional and relational communication [17, 5], online and offline modality switching [48], and evaluation of in-group and out-group members [10]. The theory has found many applications in a variety of domains beyond human conversations into computer-mediated communication (CMC), such as communicating with embodied agents [19], AI music creation [33], deviancy in online groups [47], and identification of malicious messages on Twitter [7]. EVT has shown to be an appropriate framework in describing the effect of individuals' expectations towards others in online interactions in numerous and diverse contexts – similarly, we extend these works to apply it in the context of players' responses and perceptions of their teammates' performance in online multiplayer gaming. We explain how EVT is salient in such context in the following section.

First, we explain in detail the core concepts of EVT: expectancy, violation and its valence, and the resulting responses to the violations. We then address how previous literature has used EVT to evaluate the behavior and responses of individual-to-individual interactions in group-based and collaborative settings to extend how EVT can be applied in a multiplayer online gaming situation, specifically in understanding how negative violations may lead to negative player perceptions toward their teammates in the context of *League of Legends*. Finally, we describe how the concept of "rank" in *League of Legends* holds value as a crucial interaction characteristic that sets a player's expectations towards their teammates' performance.

2.1 Key Concepts

2.1.1 Expectancy

Expectancies refer to cognitions of expectations in a given interaction. Simply put, they are the expectations that an interactant holds regarding specific individuals and their behavior. Burgoon and Walther (1990) [23] explain that such expectations are shaped by social norms, as well as three sets of interaction characteristics: the communicator, the relationship, and the context [17]. *Communicator characteristics* highlight the individual's features that are relevant to the interaction, such as their personality, identity, and appearance. *Relationship characteristics* refer to the relationship between the interacting individuals, including familiarity, status, and attraction. *Context characteristics* are environmental factors and cues, such as formality, that may be salient to the interaction behaviors [47]. These characteristics form an important basis for understanding the possible expectations that arise from them.

The expectations can be further divided into *pre-interactional* and *interactional* expectations based on when the expectations are formed and applied. Pre-interactional expectations are expectations formed prior to the interaction, based on past experiences, norms, and beliefs [20]. Meanwhile, interactional expectations are expectations that develop over the course of the interaction and are influenced by the various cues and climates of the interaction [20]. Previous literature has studied how pre-interactional expectancies cause individuals to evaluate another's behavior differently [17, 49]. Similarly, this study specifically focuses on pre-interactional expectancies set by the individual player through communicator characteristics provided by the game. In Section 2.2, we delineate in depth the specific types of expectations addressed in the scope of the study.

2.1.2 Violation and Valence

In human-to-human communication, expectations are either confirmed or violated by an individual [22]. The violation refers to a deviation from the expected behavior [16, 19]. For example, a customer suddenly shouting profanities in a store would be considered a violation of normal consumer behavior. The scope of violation has expanded beyond simple communicative violations to include emotional or relational violations [5], especially in online contexts [11, 7].

EVT introduces *valence*, or assessment of whether the violation is positive or negative, of violation as a core influence on how an individual responds to a violation. Other theoretical expectancy models have concerned *met-expectations* viewpoint of expectations, investigating the effect of confirmed expectations on satisfaction and other factors [25, 12, 15]. However, EVT expands on such work by emphasizing the importance of the valence of both the violator (*reward valence*) and their action (*violation valence*) [17]. When a violation occurs, a dual-appraisal process of "interpretations and evaluation" of the violation is triggered [18]. An individual will aim to attribute how positive or negative the violation was, with possible linkage to the individual's perceived valence of the violator.

Through this process, a behavior can be categorized into four categories according to EVT: *Positive Confirmation*, *Positive Violation*, *Negative Confirmation*, and *Negative Violation*. *Positive Confirmation* is an act that is expected and positively accepted, while *Positive Violation* is an action that exceeds one's expectations. For example, a colleague finishing their work early and exceptionally well is normally considered a positive violation. Similarly, *Negative Confirmation* is an undesirable yet expected action, while *Negative Violation* is both undesirable and unexpected [19]. An example of a negative violation is when a student keeps disrupting a lecture by making inappropriate comments.

2.1.3 Violation Response

Crucially, EVT anticipates that violations in either valence direction – positive or negative – will yield stronger effects on a person's perception and response than confirmation. Numerous studies have highlighted that positive violations will yield more favorable outcomes than positive confirmations [4, 48]. For instance, the impact of positive violations on satisfaction is expected to be stronger when perceived performance surpasses rather than meets the expected level of performance [19, 21]. Conversely, negative violations have shown to yield less preferable outcomes when compared to even negative confirmations [20, 21]. The theory claims that violations with a negative valence will have a more diminishing effect on outcomes compared to instances where expectations are confirmed.

Specifically, the theory predicts that violations are more likely to lead to more extreme social judgments compared to when expectations are met [4, 48]. These judgments manifest in various dimensions, encompassing aspects such as performance, trust, credibility, and intentions for future interactions [34, 4, 8]. In online contexts, for instance, larger infractions of negative violations on Facebook (e.g., posting inappropriate images of an individual) showed more instances of individuals deleting the content or even "unfriending" the problematic individual in extreme cases [44].

This study focuses on the *negative violation* of a player's expectation regarding their teammate's performance due to its potential connection to the within-team conflict that occurs between individuals. According to EVT, such negative violations are expected to result in greater negative arousal and more severe social judgments towards the player and their gameplay. The consequential arousal and judgments can induce the player to more aggressively respond to the "offending" player and foster conflict. Thus, the study will evaluate whether and how such negative violations occur and observe their effect on player arousal and attribution toward the teammate.

2.2 Application of Expectancy Violation Theory to *League of Legends*

We apply the framework of EVT to multiplayer team-based online multiplayer gaming to explain its potential effect on negative player arousal and behavior. We specifically approach the problem in the context of *League of Legends*, a popular Multiplayer Online Battle Arena (MOBA) that has been used to investigate and understand the within-team conflict in competitive gaming contexts [37, 36, 40]. First, we discuss the *League of Legends*' teammate interactions. Then, we note previous research that has applied EVT in a similar context as team-based competitive gaming. In conclusion, we address how the characteristics of *League of Legends* and its design lend it to players potentially relying on pre-formed expectations to judge their teammates' performance.

2.2.1 Research Context: League of Legends

League of Legends is a highly popular Multiplayer Online Battle Arena (MOBA) video game [31]. In League of Legends, players assume the role of "champions", each with unique abilities and playstyles, and form teams of five to compete against each other. The primary objective is to destroy the opposing team's Nexus, a building located in their base, while communicating with your teammates, fighting the enemy team, and competing over objectives.

A player regards their teammate under the context of their position and the corresponding champion within the game. A position in *League of Legends* refers to the main role of a player amongst five potential categories: Toplaner, Jungler, Midlaner, Attack Damage Carry (ADC), and Support. Top and Midlaners fight one-on-one with the enemy team's corresponding members in their lane, while the ADC and Support group together against the enemy team's ADC and Support duo. Lastly, Jungler roams between the lanes hunting monsters and helping other laners. A player typically selects a champion that fits into their position.

Thus, in *League of Legends*, a player interacts with their teammate in accordance with fulfilling each person's perceived "responsibilities". Each player will have certain commonly-agreed objectives or roles they perform based on their position and champion selection (e.g. *"An ADC will hang back and do massive damage"*). However, how *well* they perform these tasks will vary greatly depending on the player's skill, and thus collaboration requires the assessment of the player's performance. These assessments are naturally connected to a player's expectations of the teammate's performance.

2.2.2 Expectancy Violation Theory in Group and Collaborative Contexts

It should be noted that *League of Legends* and other similar gaming contexts are atypical of the settings in which EVT has been applied. The majority of EVT literature has focused on partnerships [45, 25, 48] or communication as its main task [19, 44]. EVT generally addresses an *individual*'s expectations towards another, single entity (e.g., another human, embodied agents [19], or AI model [52]) in an *isolated, communicative* context. Such assumptions ensure that the interaction is not affected by the introduction of intragroup or task-dependent interactions.

However, more recent research has shown EVT to be a useful model for understanding individual expectations towards another individual in group and collaborative settings. Nicholls and Rice (2017) explored responses to deviance in online groups and communities by incorporating a social identity approach at the group level, but also EVT for an individual-level perspective of how an individual reacts to deviance [47]. Lee (2021) conducted a study analyzing how prior expectations moderated the effect of tolerance and rejection on minority group emotions and needs by having the participants play Cyberball [53], a game where participants toss virtual balls at each other to study ostracization. They observed that expectations had a significant impact on social acceptance, tolerance, and other domains [39]. For expectation in collaborative settings, Joardar (2011) examined a workgroup's change in acceptance of when a newcomer from a different culture violates group expectations of performance set based on preconceived stereotypes regarding the culture. The study also further highlights the effect of pre-interactional expectations on judging a player's performance [34].

We emphasize that not only does *League of Legends* lies in a similar group-based and collaborative context as these studies, but the expectations shaped toward an individual still follow the **traditional individual-to-individual** interactions as adopted by the majority of EVT-related work. Thus, we believe that EVT can be applied to describe the expectations and violations incurred between two individual teammates in a game session that is embedded in the context of group work.

2.2.3 Expectations and Violations in *League of Legends*

To expand on the above work, we posit that *League of Legends* possess characteristics that lend it to analysis through the lens of EVT, especially zoning on expectations set by pre-interactional communicator characteristics. The two main attributes of the game that we address are (1) high interdependence between teammates necessitating expectation setting and (2) the temporary nature of the team formation limiting characteristics that shape expectations. These factors may contribute to the greater reliance on pre-formed models of their temporary teammates.

League of Legends has been noted for its high interdependence between players to achieve success, lowering the individual control over the outcome of the game [36]. This necessitates that players share a similar mental model of not only the ongoing game but also each other and their behavior to ensure success. Yet the game is also defined by its complex mechanics – the player must focus on champion skills, map status, ally and enemy placement, synergies, and many other variables – that are interwoven with other players' play. Thus, accurately expecting how each of these factors will play out is difficult to predict, especially for players at lower levels. When a player inevitably doesn't follow the player's ideas on how they are "supposed" to have played, it serves as grounds for potential conflict.

More importantly, the temporary nature of the "team identity" between players limits the information available to the players on their teammates. The latter characteristic is common in many other team-based competitive multiplayer games, such as Overwatch [29] and Valorant [32], where teammates are constantly shifting and rarely form long-lasting connections beyond the 20 to 30-minute game time they share. This impermanence of interaction between the individual team members means that players do not hold expectations specific to a certain player; rather, they must rely on game norms or other game-given factors. To each player, the interaction characteristics of the interaction in gaming may seem like they stay consistent between the games, even though the communicator may vary greatly despite the game's best efforts to match similarly performing players. Thus, we posit that pre-interactional characteristics will play a large role in forming player expectations.

2.3 Role of Rank in Player Expectations

To explore this idea, we employ "rank" as the basis for understanding and evaluating player expectations in this study. In *League of Legends*, rank is a performance metric given to players who participate in the Ranked mode of the game. There are 10 tiers (ranging from Iron to Challenger) and four sub-level divisions within each tier (not including the Master tier and beyond). The game tries to match players of similar performance based on its internal scoring system of the rank called matchmaking rank (MMR). The player then earns ranking points (named League Points (LP)) determined by the outcome of the match and their performance within the game.

We posit that rank is a crucial *communicator characteristic* that shapes a player's *pre-interactional expectations*. Unlike other online human-to-human contexts such as social networking services (SNS), games such as *League of Legends* focus on short bursts of interaction between individuals that are not meant to continue beyond the single game. Thus, the player will can only create an image of the specific player based on the limited information provided. The game provides information, such as last season's rank (provided in the form of a banner around the player's profile), champion mastery for certain characters, and honor level. Other players may choose to access third-party websites (such as OP.GG [3]) to gain a more detailed insight. However, we emphasize that rank is the all-encompassing and the highest-priority representation of the player's *performance*. Additionally, unlike other factors that the player must actively seek out to see, *Rank* is immediately visible (by the banner) or is already assumed by the player based on the matchmaking system. We propose that rank may be synonymous with other pre-interactional communicator characteristics, such as a newcomer's racial identity [34]. Thus, we anchor the research on the notion that rank contributes a large part in molding a player's expectations of their teammates.

Chapter 3. Hypotheses

Based on the previous work on EVT and its extension into the context of *League of Legends*, we formulate hypotheses on the accuracy of the players' expectations for a teammate's performance based on rank, how rank may be used to observe the occurrence of negative violations, and whether negative violations of expectation will yield more negative arousal and attribution towards the violator.

3.1 H1: The Misalignment of Expectations on Rank

As described in Section 2.3, we employ "rank" as the pre-interactional communicator characteristic of a player's expectations toward another teammate's performance. However, there is an evident lack of specific definitions and measures that surround "rank". Namely, how a player at a rank is "supposed to be" is unclear and not agreed upon. We therefore believe that the players will exhibit wide-ranged, and thus generally incorrect, predictions on what a player's rank is by observing their behavior. If true, this wide array of expectations formed on the same rank by multiple players may hinder proper teammate performance expectation setting for players and open up opportunities for conflict.

H1: Players hold inaccurate and wide-ranged expectations of rank on a player's performance.

3.2 H2: The Effect of Expectations of Rank on Negative Expectancy

Based on previous literature as described in Section 2.1.2, negative violation, individuals go through a dual-appraisal process of interpretations and evaluation of the other individual's behavior in an interaction, especially when a violation occurs. As H1 predicts that players will exhibit inaccurate observations on what rank a player should be, we predict that such expectations formed through rank may be violated when a teammate makes an action that negatively contributes to other teammates, starting this process. To check this, we observe whether a player's expectations towards their teammate are higher for those at higher ranks by showing a player their teammate's same unfavorable behavior (such as mistakes, getting killed, or failing an objective) as played by differently ranked players. We propose that showing a player's unflattering or "bad" actions at a higher rank will set the player's expectations higher for that player, thus causing the mismatch in expectedness to be greater and the action itself to be considered more negative and unexpected.

H2: Players form higher expectations to perceived higher ranks, whose violation is more unexpected and negatively perceived than for lower ranks.

3.3 H3: The Effect of Negative Violation on Player Perception

Various previous literature [4, 20, 21, 48] have demonstrated that negative violations cause greater levels of arousal in an individual, especially in a less desirable way. Regardless of how the expectation is formed, we believe that the violation of a player's expectation towards their teammate's skill level can happen when watching a player make a mistake or undergo loss. Thus, we confirm whether a negative violation of such expectations indeed results in greater levels of negatively perceived emotion and increased blame on the individual in comparison to when expectations are met in either direction.

H3: A teammate's negative violation of performance expectation will lead to more negative emotional response and attribution toward the teammate than confirmation.

Chapter 4. Methodology

This study constructed a between-subjects survey experiment to investigate the effect of expectation from ranks on expectancy violation. The players perform two tasks: in Task 1, we measure the player's expectations towards rank by having players predict the rank of a player, while in Task 2, we measure the expectancy and emotional and attribution response of the player when watching a clip under the consideration for a given rank. In this section, we explain and rationalize the design choices for the study.

4.1 Data Collection

To ensure the consistency and quality of the game clips used during the study, we collected recorded gameplay and replay files from current *League of Legends* players. All data collection was conducted over a month during August 2023 to validate their recency for the tasks between Patch 3.15 and 3.17.

4.1.1 Participants

We recruited *League of Legends* players to submit recordings and game-provided replay files of their Ranked Solo Queue games. We instructed recruits to record both the video and audio of each full gameplay over a week. The players were asked to play the game as naturally as possible, and the reason behind the collection of the data and its usage was not provided to further enhance the normalcy of the play. However, the authors obtained explicit consent from all the players on future usage of the clip in other tasks with full anonymization promised. The collection was conducted under the ethical approval from the Behavioral Research Ethics Board of the institution.

We recruited 59 players, all of whom played mainly on the Republic of Korea server (KR). The recruitment call was advertised across various university platforms and online communities for *League of Legends* on Korean social media platforms. From the 59 players recruited, a total of 38 players successfully submitted eligible clips. The participants ranged between 19 and 33 years old (mean=24.1, SD=3.8) and were primarily male (male=36, female=2), likely due to the dominance of young male players in the South Korean player base.

4.1.2 Game Selection

We limited the games specifically to *Ranked Solo Queue* mode, as the study design uses rank data as part of its experimental design. In ranked mode, the game places the player in a pool of similarly performing players based on their rank and recent game result history. The Solo Queue option pits the player against strangers who they never or very rarely encountered. An average ranked game of *League of Legends* lasts from 20 to 30 minutes, though players may choose to forfeit the game early or leave in the middle of the game with a penalty. We only accepted games longer than 20 minutes to ensure that the games provided meaningful interactions between players. Through these criteria, an initial set of games was collected (N=360).

We then filtered the data for players who placed in the Bronze, Silver, and Gold ranks. We choose these tiers as players in these three tiers make up 58% of the player base [1] and have a general understanding of the game that the players at the lowest rank (Iron) may not. We do not observe higher ranks due to the difficulties in gathering sufficient sample size and the belief that the regional differences between servers may be greater at higher ranks. We further limited the videos to those of players who have stayed in the same rank tier from when the data was collected to when the study was conducted (August to December) to ensure that the rank reflected the player's actual level of performance. We note that *League of Legends* season for Ranked mode typically ends in early January, meaning that the ranks in December are a close representation of the player's actual rank during that season. These resulted in 73 valid videos. Finally, we filtered the data for completed games and for which valid replay files were submitted (N=38). We then applied additional criteria for each task, which are outlined in Sections 4.2.2 and 4.2.3.

4.2 Procedure

The online study was deployed on Prolific [2], an online research platform that helps researchers recruit participants. Participants were paid 6 GBP (=7.62 USD at the time of the study) for their participation, which took less than an hour to complete. Participants first answered questions about their demographics and gaming history and then completed two tasks. The task design and procedure are outlined below. The study was conducted under the ethical approval from the Behavioral Research Ethics Board of the institution.

4.2.1 Participants

Through a set of questions, we only selected participants who have played Ranked Solo Queue mode during the current season (Season 2023). Participants were required to correctly answer questions that required knowledge of the game to further ensure that the participant was a *League of Legends* player. The participants answered a questionnaire on their last season and current season rank and how long they have been playing the game, excluding extended breaks. The players were also asked to rate generally how well they think they play in comparison to their teammates as well as how well they think their team plays in comparison to the enemy team. A total of 539 participants completed the study. After limiting to only Bronze, Silver, or Gold rank players, 106 participants remained. We further eliminated participants who failed attention checks, had inconsistent answers, or did not complete the survey properly. A final total of 80 participants remained. Table 4.1 outlines the demographic information and distribution of the participants based on server and rank.

4.2.2 Task 1: Investigating Expectation Alignment of Ranks

Task Design

To investigate H1, the first task provided a 9-minute video clip of a player whose rank information was not provided and asked the participant to guess the player's rank tier. The participant was also asked to explain the reason behind their guess and rate how confident they were in their answer on a 5-point Likert scale. The player was always assigned a video at their rank (i.e. a Gold player to a Gold-rank game). This was done as players may not be familiar with other ranks, especially at lower ranks, and as players will have greater and guaranteed exposure to players at their tier. Thus, we intended to maximize the chance that the participant can guess correctly.

Table 4.1: Demographic information of the study participants

Variable		Total Sample (n = 80)
Gender (n, %)	Male	61 (76.25)
	Female	18 (22.5)
	Non-binary/Non-conforming	1 (1.25)
Server (n, %)	Europe Nordic and East (EUNE)	12 (15.0)
	Europe West (EUW)	28 (35.0)
	Japan (JP)	1 (1.25)
	Latin America North (LAN)	7 (8.75)
	Latin America South (LAS)	6 (7.5)
	North America (NA)	20 (25.0)
	Oceania (OCE)	3 (3.75)
	Russia (RU)	1 (1.25)
	Turkey (TR)	2 (2.5)
Rank (n, %)	Bronze	8 (10.0)
	Silver	31 (38.75)
	Gold	41 (51.25)
Play frequency (n, %)	Every day	3 (3.75)
	Several times a week	35 (43.75)
	Once or twice every week	30 (37.5)
	Once or twice every month	12 (15.0)
Average play session duration (# of games)		5.85
Preferred roles	Top	24
	Jungle	21
	Middle	26
	ADC	14
	Support	39
	No preference	11

Game and Clip Selection Criteria

We selected two videos for each rank tier for a total of six videos. Several criteria were used to maximize the representativeness and diversity of the videos from the dataset in Section 4.1.2. First, the game must have been completed without intentional misplays or sabotage from the player. Second, the player must not be the best or worst player on the team, evaluated through a third-party website that analyzes the game results [3]. Third, the champion that the player plays must have been released for more than a year to reduce unfamiliarity and must be agreed upon that their mechanics are generally understood among the player base. Fourth, the player must have shown plays that fit the general rank norm. Finally, various positions and champions were selected across the videos. One of the authors with expertise in the game analyzed the videos and selected games that met the criteria, and a second author confirmed that the videos fit these criteria. For each rank, one video depicted a game in which the player's team won, while the other video showed a game in which they lost. We intended to avoid potential bias in the guess results due to the game state. The videos were randomly assigned to each participant.

To provide robust information, the video consisted of three segments of 3 minutes from the beginning, middle, and end portions of the game. Each three-minute segment was selected from the middle of each portion, which was divided equally into three based on the game length. The time length was chosen to provide sufficient time and context across the whole game. The video provided information as if the player was watching a teammate in the actual game by limiting the information provided in the game. However, due to the limitations of the replay file from which the video was recorded, the players also saw the overall game score. The video control features, such as playback speed or timeline, were disabled to prevent players from skipping or skimming through the content. To check that the participant fully watched the video, we inserted photos of animals in the middle of the video and asked the participants to select the photos that were shown during the video. We eliminated any participant who did not correctly answer the question. Detailed information on the videos for Task 1 is provided in Table 4.2.

4.2.3 Task 2: Investigating Effect of Negative Expectancy on and Player Perception

Task Design

For the second task, the participant was provided with a 2-3 minute clip of a player with a label of their "rank tier". The study consisted of two conditions: Condition 1 showed the rank to be a rank lower than the actual rank (e.g. "Silver" for a Gold-rank player's game) and Condition 2 showed the rank to be a rank higher than the actual rank (e.g. "Platinum" for a Gold-rank player's game). In reality, the rank of the player was the same as the player watching the video clip. The disparity between the two conditions was made to observe the effect of rank on the player's expectation.

For the first part of the task intended to answer H2, the player answered a questionnaire on how the player behaved according to their expectations (Burgoon and Walther's expectancy scale) [19]. They were then asked to answer how well they thought the player performed, what they thought the player did wrong, and how well they thought that the rank matched their play and why. For the second part, the participants answered a questionnaire on their emotional response to the clip shown (Game-adapted PANAS-like scale [50]) and how they attributed the blame of the player's actions during the gameplay (Game-Specific Attribution Questionnaire (GSAQ) [28]). Finally, the players answered questions that asked them to further explain the reasoning behind their answers ("*For the emotions you ranked high, why do you think you would feel such emotion?*").

Game and Clip Selection Criteria

We selected one video for each rank tier for a total of three game clips for this task. For this task, clips depicting a player making "bad play" that resulted in them being killed by enemy players were chosen to observe *negative* violations. The same criteria from Task 1 were applied to the videos. Additionally, one of the authors with expertise and experience in the game went through the collected data and chose clips that displayed an instance of a misplay by the player, was not an "egregiously" bad performance, and whose actions did not seem intentional. The choice was then validated by another author with knowledge about the game to ensure that the clip did not display outlier behavior from the rank. We also ensured that the clip did not provide context of the end game state by selecting clips from the early to mid-stages of the game and that the games were not skewed towards one team.

The video length was a minimum of 2 minutes to include sufficient context of the situation and for the participant to assess the player more accurately without introducing external effects from other variables. The video controls were disabled. The same attention check method from Task 1 was used to verify the validity of the answers. We eliminated any participant who did not correctly answer the question. Detailed information on the videos is provided in Table 4.3.

Measures

The second task measured three scales. To answer H2, we adapted Burgoon and Walther's 10-item measure of expectedness and evaluation using a 7-point Likert scale (1=unexpected or negative valence; 7=expected or positive valence, respectively) [19]. The measure included items such as "*My teammate behaved differently than I anticipated.*" and "*My teammate's way of playing was undesirable.*", whose wording had been modified to better fit the gaming context. For H3, we show a list of different words that describe feelings and emotions and ask the participant to rate how much they are feeling each emotion on a 7-point Likert scale [50]. Specifically, we use words that have been applied to the gaming context, such as "Malicious Delight". We also use a Game-Specific Attribution Questionnaire (GSAQ) developed by Mandryk et al. (2017) [28], which adopts previous attribution questionnaires that measure how the participant attributes the cause of one's behavior or actions along different dimensionalities to the gaming context. GSAQ contains four dimensions: internality (4 items), stability (3 items), controllability (3 items), and globality (3 items). However, we eliminated globality, which measures whether the cause applies across multiple situations as it did not apply to our game-only context, creating an 8-item instrument. We also further removed questions that did not pertain to the context of attributing the cause for specifically the teammate's actions, for one item from internality and controllability each, resulting in 8 items total. We further altered the wording of the GSAQ to better measure the attribution of a player's performance in a team-based game context.

Table 4.2: Video selection for Task 1

Condition	Rank (August)	Rank (December)	Length	Champion	Position	# of subjects
Bronze (Win)	Bronze 4	Bronze 4	24m 13s	Rammus	Jungle	2
Bronze (Loss)	Bronze 3	Bronze 3	37m 15s	Akshan	Mid	6
Silver (Win)	Silver 3	Silver 2	24m 42s	Neeko	Mid	13
Silver (Loss)	Silver 3	Silver 4	20m 48s	Karma	Top	18
Gold (Win)	Gold 3	Gold 4	28m 30s	Alistar	Support	13
Gold (Loss)	Gold 4	Gold 4	28m 50s	Wukong	Jungle	28

Table 4.3: Video selection for Task 2

Condition	Outcome	Version	Clip Length	Champion	Position
Bronze	Win	13.15	2m 5s	Vayne	Top
Silver	Loss	13.17	2m 22s	Nautilus	Support
Gold	Win	13.17	2m 6s	Kennen	Top

Chapter 5. Results

We present the statistical and qualitative results for each hypothesis. We also discuss the analysis of the results and their potential causes briefly, with a more in-depth analysis in the Discussion section. For all statistical comparisons unless stated otherwise, the Kruskal-Wallis test was chosen as the statistical comparison method since the data collected was ordinal data, such as rank prediction or Likert scale answers and we could not make assumptions about the normal distribution of the answers.

5.1 Rank Prediction

We first checked the distribution of the rank predictions at each rank to observe how accurately players were able to match the rank based on performance. The distribution of each rank for all players of each rank can be seen in Figure 5.1, with the median rank outline in blue. We observe an overall trend of players over-predicting the rank of the player shown in the clip for Bronze and Silver ranks, as the median is shown to be Silver and Gold for Bronze and Silver, respectively. However, it should be noted that the sample size for Bronze is quite small, potentially making the distribution not representative of the whole population. Only the gold rank shows a median value that matches the actual rank of the player in the video. Even so, we note the wide discrepancy between players' answers for Gold rank. We also calculated the Interquartile Range (IQR) for each rank: 1.25 for Bronze, 1.0 for Silver, and 2.0 for Gold. The median and IQR values demonstrate an overall large variance of answers among the players in both directions (Table 5.4).

In looking at the distributions, we observed that the players who watched the game in which the player lost seemed to have generally lower distribution for those who watched the game in which the player won. To confirm whether the game state affects the participant's assessment of the player's rank, we performed a Kruskal-Wallis test (α=0.05) of the distributions of rank tier predictions for each rank based on the game result of the video (win or loss) for Silver and Gold rank (Figure 5.2). We eliminated the Bronze rank as the number of samples was too small to perform a statistically meaningful comparison test. The p-values across both ranks were large ($p=0.772$, $p=0.627$), signifying that the differences were not statistically significant. However, the median of the two conditions differs for both ranks of a difference of one and two ranks, respectively, showing that the game state and the end result may affect how players estimate the rank of a player.

Checking the average of participant's confidence level for their answers across the three ranks (3.13 for Bronze ($SD=0.64$), 3.29 for Silver ($SD=0.54$), and 3.76 for Gold ($SD=0.69$)) shows that players are quite confident in their answers (Figure 5.3). Only four players answered 3 or below for the question, showing that all players were at least neutral or above on how confident they were in their answers. We also see a general trend of higher confidence as the rank of the player increases.

Based on the results of Task 1, we observe that participants have a highly incongruous spread of ranks despite being provided with the very rank they belong to. While some of the spread may be attributed to the differences in rank division or skill level between regions, the large disparity in the rank predictions shows that the rank expectations between players are misaligned and inaccurate. Thus, H1 is confirmed.

Table 5.1: Interquartile range and median of predictions for each rank

	IQR	Median
Bronze	1.25	Silver
Silver	1.00	Gold
Gold	2.00	Gold

5.1.1 Perfomance Prediction for Teammates

We also check the consensus on how players tend to view their teammates with respect to their own performance as well as with respect to their opponents. Based on the 5-point Likert-like scale with polarity (1 - Much worse than me/Much worse than my team, 5 - Much better than me/Much better than my team), the participants rated how well their teammates with respect to the participant and the opposing team. The mean for participant-teammate comparison was 2.9625 (SD=1.04), and 3.225 (SD=0.886) for ally team-enemy team comparison. Thus, we observe that players tend to rank their teammates slightly lower or at the same level as them, but seem to be rating their opponents to be better at the game than their teammates.

We conducted a Kruskal-Wallis test (α=0.05) to ensure that there was a significant difference between the two distributions. However, with a p-value of 0.0597, we were unable to reject the null hypothesis. Though we do not see a statistically significant test result, it is an interesting viewpoint of how players seem to view the enemy team to be better, despite being put in a randomly drawn pool of players to form the temporary groups. This type of disproportionate belief in their teammates' performance may be a resident part of a player's assumption.

5.1.2 Qualitative Analysis

To understand in detail how players formed rank expectations, we analyzed the answers to why the participant guessed the rank. Two of the authors conducted open coding of the question responses regarding the reason behind the participant's rank.

Sentiment of the Performance Perception

Among the respondents, we noticed a distinct trend of those who brought up positive, negative, or both aspects of the play to reason their prediction. The distribution between the positive and negative-focused comments was similar (27 and 25, respectively). However, all participants who focused only on the positive aspects of the teammate's play had over-predicted the rank. Meanwhile, those who had focused on the negative aspects or mentioned both aspects had mixed predictions. This shows that players who hold a negative view of their teammates have a more inconsistent assessment between the players.

Reliance on Subjective Factors

Many of the players described various parts of the gameplay to reason their decisions. Analyzing what specific standards they used reveals that the factors used to assess rank were overwhelmingly factors that relied on a player's subjective understanding of a "skill", such as macro abilities, map awareness, push opportunity, and more. For example, P15 mentioned that "*[The player] had some skills on how to* *farm jungle, but not so much on...macro play and ganking creativity and effectiveness. Also, some fights were obvious they wouldn't go well and he still went for it.*" These are performances that are difficult to quantify and require that the observer also has an accurate assessment of how such skills must be carried out. The focus on such more long-term, macro-level performance seems to show that players are relying more on an overall assessment rather than specific visible measures.

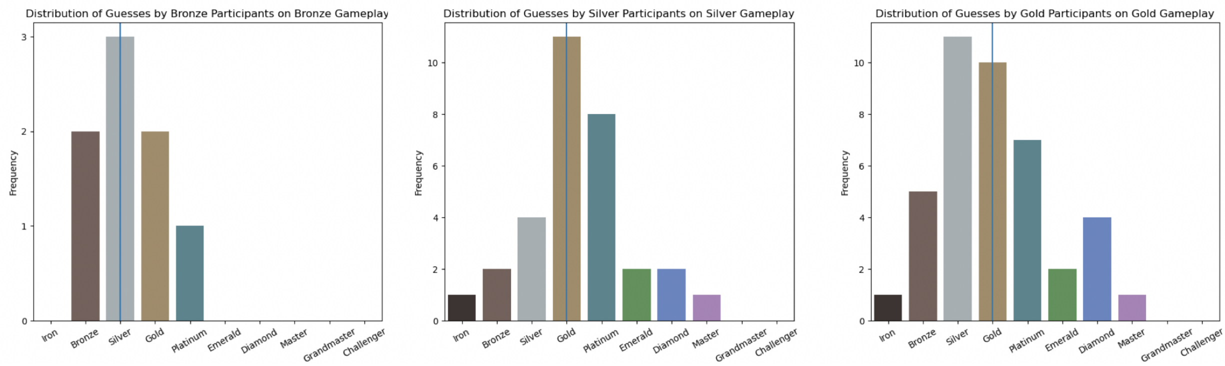

Figure 5.1: Distribution of rank predictions from the players watching the gameplay of the same rank as themselves (left: Bronze, middle: Silver, right: Gold). The blue line shows the median rank of the distribution. Both clip conditions were combined.

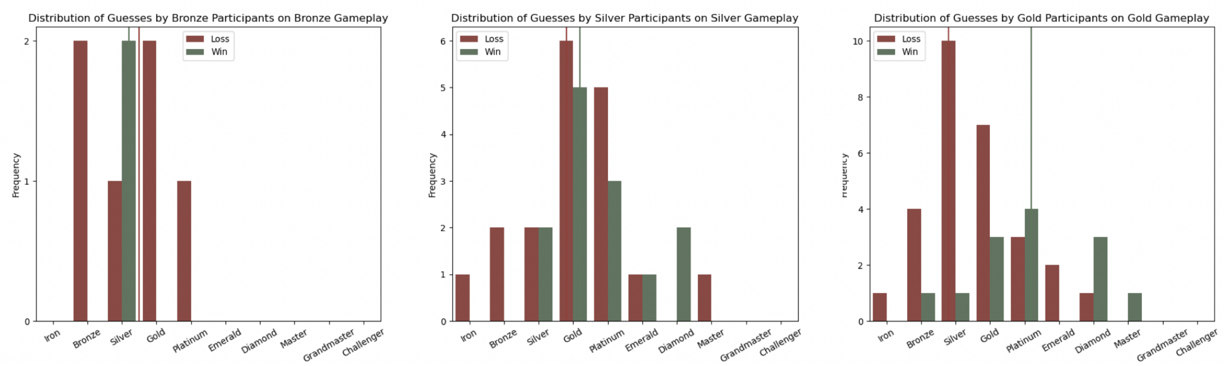

Figure 5.2: Distribution of rank predictions from the players watching the gameplay of the same rank as themselves (left: Bronze, middle: Silver, right: Gold). The blue line shows the median rank of the distribution. Red denotes the video in which the player's team lost, while Green denotes the video in which the player's team won.

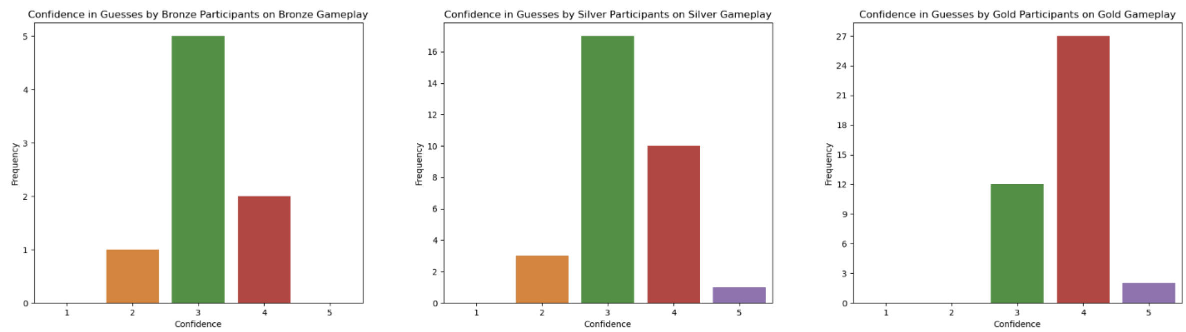

Figure 5.3: Distribution of confidence level from the players on their rank prediction.

5.2 Effect of Rank on Expectancy

We then compared the expectedness and evaluation values across the two conditions for Task 2 (Condition 1: lower rank labeled, Condition 2: higher rank labeled). The participant distribution for each condition is shown in Table 5.2. To test this, we performed a Kruskal-Wallis test (α=0.05) of the expectedness (averaged over five items of expectedness) and valence (averaged over five items of valence) based on the condition for rank shown (high or low). However, no significance was found between the two distributions across both expectedness and evaluation (Table 5.3).

In case participants showed a consistent trend of over or underguessing the ranks of the player shown, we also divided the groups based on whether the participant had guessed higher or lower than the actual rank from Task 1. We hoped to validate whether accounting for how users tend to direct their expectations made from ranks may make the variables more controlled. Eliminating the 11 players who had accurately predicted rank, 64 participants were the final sample size for this test, with 44 for those who had predicted higher than the ground truth and 20 for those who had predicted lower. Thus, we conducted a 2x2 Kruskal-Wallis test (α=0.05) along the two conditions, yet still did not find a statistical difference between any of the conditions.

If H2 holds, we expect to see a statistically significant difference between the two conditions, in which one condition results in a higher expectedness and/or valence. However, we once again did not see any correlation between any of the conditions given. Thus, we fail to reject the null hypothesis for H2.

5.2.1 Qualitative Analysis

To understand in detail why we did not observe expectancy differences based on the rank shown, we analyzed the answers to whether the player believed that the rank shown was correct and why so.

Acceptance of the Rank Shown

The majority (57.5%) of the participants replied that they thought that the rank they saw was accurate to the player's abilities, regardless of the condition provided. It should be noted that most of the respondents who stated that rank should be lower had been subjected to Condition 2, showing that their players' evaluation still holds some valid comparison. However, very few respondents stated that the rank should be higher, even when shown the video at a lower rank, showing that players seemed to perform an inaccurate assessment of the player's performance level when bad performance is shown.

Vague Assessment Metrics

The participants provided general and vague information on why they believed the rank was aligned or not, often describing the players as a "typical X rank player" or stating that the play simply matched what a certain rank is like. A few players mentioned more specific measures, such as "consistently last hitting and...keeping the lane warded" (P78). The results hearken back to the answers from Task 1, which again focuses on more subjective standards, such as the player's decision-making skills. However, it should be noted that players also gave leeway on the fact that the player may have been having a bad game ("*It does not match their rank at all, but maybe I am just seeing a very unfortunate time for this player.*", P13) or may have not been familiar with the champion ("*I mean, maybe they don't usually play this role or champion, but no I don't think with that gameplay that gold is reflective of their rank.*", P38). Players also sometimes made comparisons to their own skill level as the basis of performance comparison.

Table 5.2: Participation distribution in Task 2 for each condition by expectation direction from Task 1

Expectation direction	Condition 1	Condition 2
All	46	34
Over-predicted	21	23
Under-predicted	11	9

Table 5.3: Kruskal-Wallis test results for expectancy for each expectation direction group

	Expectancy		Valence	
Expected direction	Chi-square	p	Chi-square	p
All	0.286	0.335	1.946	0.163
Over-predicted	1.4925	0.2218	0.0311	0.86
Under-predicted	1.000	0.317	0.540	0.462

5.3 Negative Violation and Player Perception

To investigate whether negative expectancy violation will negatively affect a player's perception of and reaction to their teammate (H3), we first categorize the data points into the four categories of expectancy along the median line. The four quadrants are defined as Positive Confirmation (PC) if expectedness is low and valence is also high, Positive Violation (PV) if expectedness is high but valence is high, Negative Confirmation (NC) if expectedness is low and valence is low, and Negative Violation (NV) if expectedness is low and the valence is also low. We show the scatterplot of the distribution in Figure 5.4 and display the number of samples for each expectancy in Table 5.4.

Though NV does make up the biggest part of the sample size, surprisingly, we see the existence of numerous PC and PV samples despite the intentional selection of clips that exhibit negative outcomes and missteps of the players in them. These participants seem to have viewed the clip positively regardless of the failure displayed in the video, suggesting that they may have regarded the play as just a natural part of how the game proceeds, failed to grasp the severity of the missteps or saw positive aspects of the play despite the bad result. Overall, this could mean that player deaths or damage is not always considered unfavorable by others, further murking the understanding of how a player may be reacting to different plays.

We first test H3 on whether NV is correlated with higher levels of negative emotion and greater attribution in comparison with when the behavior is expected. As shown in Table 5.5, There is statistical significance for emotions such as Happiness, Amusement, Excitement, Anger, Relief, Frustration, and Disgust. From these, Anger, Frustration, and Disgust are all strongly negative sentiments, specifically of feeling being wronged by the other party. Thus, we see that negative violation in gaming does lead to an emergence of higher negative feelings for the observer, ultimately opening up a potential pathway to greater conflict within the game.

To further check the direction of the correlation as well as the differences between the confirmations based on valence rather than just between expected action and NV, we performed a posthoc Dunn Test with Bonferonni correction (α=0.05). We observe that Anger, Relief, Frustration, and Disgust show strong statistical differences in NV-PC relationships (Table 5.6). We also see significant differences between NC and NV for Anger and Frustration, suggesting that a negative violation, even when compared to a negative confirmation, resulted in greater negative arousal.

Finally, we tested whether the GSAQ showed any significance along the three divisions of internality, stability, and controllability using the same statistical test as before, but saw no significance.

Thus, we are partially able to reject the null hypothesis for H3 in terms of emotional response, but not for attribution.

Table 5.4: Count of sample points of each expectancy

Expectancy	Positive Violation	Positive Confirmation	Negative Violation	Negative Confirmation
Count	6	16	22	15

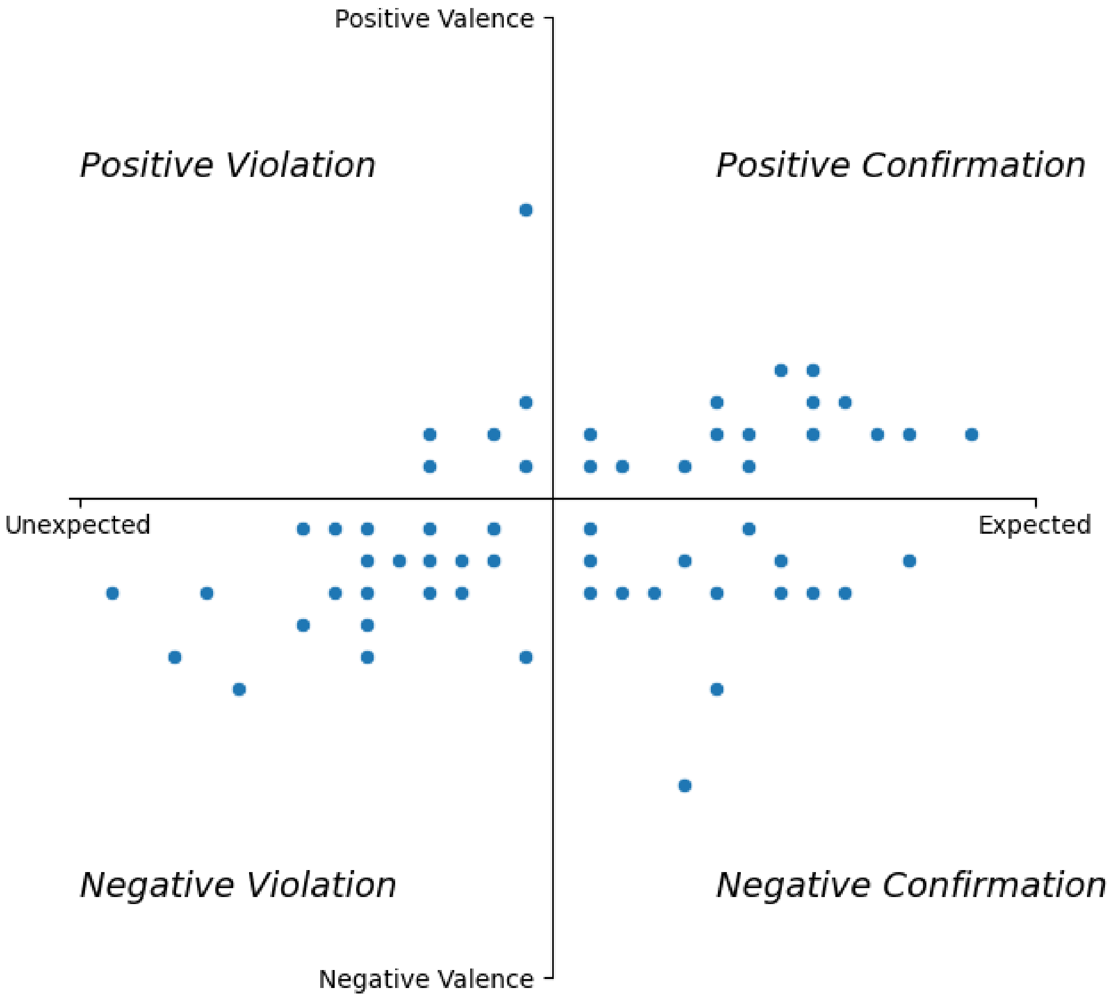

Figure 5.4: Quadrant plot of the expectancy distribution. Unexpected behavior results in violation, while expected behavior leads to confirmation. Valence denotes the direction in which expectations were violated or confirmed.

Table 5.5: Kruskal-Wallis test results for expectancy and emotion

	Chi-squared	adj. P
Happiness	**7.45**	**0.024**
Boredom	2.18	0.336
Amusement	**6.70**	**0.035**
Surprise	3.99	0.136
Malicious Delight	1.22	0.542
Excitement	**10.00**	**0.007**
Fear	0.40	0.818
Anger	**12.80**	**0.002**
Relief	**6.33**	**0.042**
Frustration	**12.88**	**0.002**
Wonderment	1.47	0.479
Disgust	**11.20**	**0.004**

Table 5.6: Posthoc Dunn Test Results for Negative Violation and Confirmation for emotion

Emotion	NC-NV	NC-PC	NV-PC
Happiness	2.02	-0.821	**-2.69**
	0.130	1.00	**0.0211**
Amusement	1.90	-0.797	**-2.56**
	0.172	1.00	**0.0315**
Excitement	2.19	-1.15	**-3.15**
	0.0849	0.748	**4.88e-3**
Anger	**-2.65**	1.08	**3.53**
	0.0243	0.846	**1.23e-3**
Relief	1.90	-0.697	**-2.47**
	0.171	1.00	**0.0401**
Frustration	**-2.73**	0.969	**3.52**
	0.0189	0.997	**1.27e-3**
Disgust	-2.28	1.28	**3.34**
	0.0682	0.607	**2.51e-3**

Chapter 6. Discussion

6.1 Lack of Consistent Expectations for Rank

The results of Task 1 show that players hold vastly different mental models of what a rank "should be". Not only are the players more inaccurate in the guess, but the discrepancy in how wrong it is goes far beyond what would be considered acceptable under the framing of "rank" as an absolute telltale of a player's performance. It brings into question how much of the inconsistency is due to the lack of proper understanding by the players, or if such understanding is achievable at all by the majority of the player base in a dynamic and complex environment such as gaming. Thus, it is important to find a set of potential explainable markers that will provide more consistent and narrowly varying expectations for players. Being able to quantify player's assessment of others' players into more realizable metrics may be the next step in reducing violations.

6.2 Unrealiblilty of Rank as an Expectation Setter

The lack of correlation between rank and expectation for H2 questions whether rank truly determines how a player views each individual. It may be possible that the expectations set by rank are too variable and random, without insurance that players hold expectations at a steady pace. Of course, we also cannot ignore that the lack of correlation may have originated from faulty study design or misunderstanding of the questionnaire prompts; but the results for H3 suggest that the measure of expectedness and evaluation seemed to accurately reflect the person's expectancy according to the strong correlation between negative emotions and NV. Namely, it is likely that rank is unreliable as an expectation setter even within an individual – a player's expectations regarding the rank may be unstable depending on the context. Thus, we may need to further investigate how a rank's expectation varies within a single player but more directly outlining their expectations.

6.3 Impact of Negative Violations in Online Gaming

Negative violations in multiplayer online gaming can occur in many ways, from how a player performs to how uncivil or uncooperative they behave. The results of H3 highlight the very real impact in shaping how a player can feel and perceive someone else's actions, especially in dealing with negative emotions that tend to redirect their valence onto others – in gaming, it often happens to be the very teammate that caused it. It may even continue beyond the singular instance, with a phenomenon known as "tilting", which carries the impact of such violations further by further skewing the expectations of the players. Thus, we draw attention to helping players maintain appropriate expectations to prevent such violations from occurring.

Additionally, we must further deepen our understanding of *why* such negative violations occur more in detail, as H2 has shown that rank alone cannot explain the phenomenon. More specifically, future research should explore each of the four expectancy categories (*Positive Confirmation*, *Positive Violation*, *Negative Confirmation*, and *Negative Violation*) to understand through specific player experiences why and when such expectancies are met. Looking at the individual in-game interactions and deriving the decision choices or contexts that players face under each condition could show the undiscovered causes of when a conflict–or does not–occurs. Such investigation can also further reason game design and its impact on diminishing or fostering team collaboration.

6.4 Implications for Platforms and Players

The results of the studies highlight significant problems in designing a "ranking" system within multiplayer gaming. Ranking is a necessary and even beneficial part of encouraging and motivating players but can cause the game to turn into a bed for greater conflict and toxicity. However, it can in turn further perpetuate the restrictive and definitive separation of the players by their rank and thus spur players to incorporate their own idea of how their teammates should perform across the games, despite the wide variance of play.

Thus, platforms may want to find more suitable metrics and methods to help players distinguish players' performance beyond simply providing statistics and rather introducing the potential meaning and interpretation behind such statistics. This may be achieved by more clearly outlining the differences between the various "ranks" or "standards" as provided by the platform, or more explicitly outlining what the players can improve and how it can be achieved rather than leaving the individuals to create a folk theory that best matches their experience. It is especially imperative at the lower levels of play that players are given more explainable metrics.

The focus of such new designs should not aim to completely fix or even lower the expectations of players — just reducing the variance between and within each player of what a certain rank or measure entails for their performance would be helpful. It could especially be beneficial if it could evolve during the game along with the evolving perception and expectations of the user, adapting to their changing mental model.

Chapter 7. Limitations and Future Work

One limitation of this study is the overall sample size, especially considering the number of conditions tested during the two video tasks. For instance, there were only eight participants whose current rank this season was Bronze. Splitting this group into the win/loss groups necessary for Task 1 made it difficult to extract meaningful information from the results for both tasks. The inherent distribution of ranks in *League of Legends*, as pursued by the game designers, also acts as a restriction on obtaining a large sample size for all ranks, as the top tier ranks (Challenger, Grandmaster, and Master) comprise at most 1% of the entire population. Due to the distribution of ranks amongst the participants, we were forced to limit the current ranks of players from Bronze to Gold. Such players mostly represent the bottom half of players in terms of skill level out of the entire *League of Legends* population. We presume that the understanding of the game that these players will tend to be worse than players from Platinum and above, which could have resulted in the large differences in expectations we saw in our results. Therefore, expanding our survey to those of higher rank, who are more likely to understand differences in player behavior across ranks, may yield different results.

Furthermore, the temporal and physical limitations of the online survey may have negatively impacted results. Participants were not given all possible game information for both Tasks 1 and 2. Notably, participants were not able to see the scoreboard of the game, which indicates important player statistics that may have helped participants make more accurate guesses. Video controls were also disabled, which meant that participants could only see important in-game events once.

Region-based differences could have affected results as well. Video clip data was obtained from gameplay footage of players on the Korean server, which was not represented at all by the participants. Instead, most players were from Europe or the Americas. A difference in skill within the same rank across servers is often mentioned within online spaces. Although the statement is controversial, it is intuitive that a region with more players and a stronger gaming culture would be more competitive than other regions. This disparity may have skewed the data in Task 1 higher, possibly interfering with any effect the conditions had on the participants.

Additionally, we may have provided insufficient experimental induction for expectancy difference for rank. Though we aimed to widen the gap between the two conditions for Task 2, the gap between the two ranks may have been too small to observe the intended effect.

Future work should not only address such experimental improvements but also take a deeper look into a more qualitative and experience-based narrative of how expectancies may be met or violated by players. Understanding from a player's viewpoint of how such problems occur will lend a useful tool in leveraging the game design to better prevent team conflict by adjusting their expectations to an appropriate standard.

Chapter 8. Conclusion

In this dissertation, we explore player attitudes towards their teammates in *League of Legends* through the lens of Expectancy Violation Theory (EVT). We theorized that players would have a mental model of expected player behavior with respect to rank, and that negative violations in which players deemed others to be of lower perceived skill compared to their actual rank would correlate with an increase in hostility and negative attitude. Therefore, we decided to test three hypotheses that cover the original theory in totality: 1) the mental models of players with regard to their expectations of player behavior based on rank are inaccurate, 2) the greater the difference between a player's actual rank and the perceived skill level of their behavior, the stronger the resulting negative unexpectedness and valence will be for the viewer, and 3) a negative violation resulting from a player's behavior leads to a greater emotional response when compared to other results.

A between-subjects survey study was held to test these hypotheses. The study consisted of two tasks, one in which participants were to guess the rank of a player of the same rank given a 9-minute clip of their gameplay, and one in which participants were asked about expectancy and emotions when watching an approximately 2-minute clip of the gameplay of a player of the same rank, but were told that the player belonged to either a tier higher or a tier lower than their actual rank. The first task effectively tested the first hypothesis; most users were unable to correctly guess the rank of the player they viewed, with some participants showing a vast discrepancy of several tiers between the player's actual rank and their perceived skill level. The second task was used to analyze the other two hypotheses; it was presumed that participants would be harsher towards the clip labeled as gameplay from a higher rank than the clip labeled as gameplay from a lower rank as differences in expectancy would be greater for the former, and that participants would consequently also show stronger negative emotion towards the former. However, we failed to reject the null hypothesis for H2 and that of attribution for H3.

Despite this, the dissertation still arrives at the meaningful conclusion that games must account for player expectations when designing a "ranking" system, as the mental models of each player vastly differ, even within the same rank. Such differences can result in greater conflict and toxicity, a bane for online multiplayer games that require strong teamwork and communication. Furthermore, with several limitations to our study, including limited sample size and potential interference from region-based differences, we believe that future additional studies may be able to show stronger evidence for the original hypotheses. Overall, this dissertation takes a stride toward understanding player attitudes towards teammates through EVT.

Bibliography

[1] League of graphs.

[2] Prolific, 2006.

[3] Op.gg, 2012.

[4] W. A. Afifi and J. K. Burgoon. The impact of violations on uncertainty and the consequences for attractiveness. *Human communication research*, 26(2):203–233, 2000.

[5] W. A. Afifi and S. Metts. Characteristics and consequences of expectation violations in close relationships. *Journal of Social and Personal Relationships*, 15(3):365–392, 1998.

[6] F. P. Alliance and A.-D. League. Disruption and harms in online gaming framework, 2020.

[7] K. Beck. Analyzing tweets to identify malicious messages. In *2011 IEEE INTERNATIONAL CONFERENCE ON ELECTRO/INFORMATION TECHNOLOGY*, pages 1–5, 2011.

[8] B. Bengtsson, J. K. Burgoon, C. Cederberg, J. Bonito, and M. Lundeberg. The impact of anthropomorphic interfaces on influence understanding, and credibility. In *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences. 1999. HICSS-32. Abstracts and CD-ROM of Full Papers*, pages 15–pp. IEEE, 1999.

[9] N. A. Beres, J. Frommel, E. Reid, R. L. Mandryk, and M. Klarkowski. 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*, pages 1–15, 2021.

[10] B. A. Bettencourt, K. E. Dill, S. A. Greathouse, K. Charlton, and A. Mulholland. Evaluations of ingroup and outgroup members: The role of category-based expectancy violation. *Journal of experimental social psychology*, 33(3):244–275, 1997.

[11] J. L. Bevan, P.-C. Ang, and J. B. Fearns. Being unfriended on facebook: An application of expectancy violation theory. *Computers in Human Behavior*, 33:171–178, 2014.

[12] A. Bhattacherjee. Understanding information systems continuance: An expectation-confirmation model. *MIS quarterly*, pages 351–370, 2001.

[13] A. J. Bisberg, J. Jiang, Y. Zeng, E. Chen, and E. Ferrara. The gift that keeps on giving: Generosity is contagious in multiplayer online games. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2):1–22, 2022.

[14] V. Blake. League of legends outlines new measures to improve player behavior, Mar 2022.

[15] S. A. Brown, V. Venkatesh, J. Kuruzovich, and A. P. Massey. Expectation confirmation: An examination of three competing models. *Organizational Behavior and Human Decision Processes*, 105(1):52–66, 2008.

[16] J. K. Burgoon. A communication model of personal space violations: Explication and an initial test. *Human communication research*, 4(2):129–142, 1978.

[17] J. K. Burgoon. Interpersonal expectations, expectancy violations, and emotional communication. *Journal of language and social psychology*, 12(1-2):30–48, 1993.

[18] J. K. Burgoon. Expectancy violations theory. *The international encyclopedia of interpersonal communication*, pages 1–9, 2015.

[19] J. K. Burgoon, J. A. Bonito, P. B. Lowry, S. L. Humpherys, G. D. Moody, J. E. Gaskin, and J. S. Giboney. Application of expectancy violations theory to communication with and judgments about embodied agents during a decision-making task. *International Journal of Human-Computer Studies*, 91:24–36, 2016.

[20] J. K. Burgoon and J. L. Hale. Nonverbal expectancy violations: Model elaboration and application to immediacy behaviors. *Communications Monographs*, 55(1):58–79, 1988.

[21] J. K. Burgoon and S. B. Jones. Toward a theory of personal space expectations and their violations. *Human communication research*, 2(2):131–146, 1976.

[22] J. K. Burgoon, D. A. Newton, J. B. Walther, and E. J. Baesler. Nonverbal expectancy violations and conversational involvement. *Journal of Nonverbal Behavior*, 13:97–119, 1989.

[23] J. K. Burgoon and J. B. Walther. Nonverbal expectancies and the evaluative consequences of violations. *Human Communication Research*, 17(2):232–265, 1990.

[24] J. Cheng, M. Bernstein, C. Danescu-Niculescu-Mizil, and J. Leskovec. Anyone can become a troll: Causes of trolling behavior in online discussions. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*, pages 1217–1230, 2017.

[25] G. A. Churchill Jr and C. Surprenant. An investigation into the determinants of customer satisfaction. *Journal of marketing research*, 19(4):491–504, 1982.

[26] C. L. Cook. Between a troll and a hard place: The demand framework's answer to one of gaming's biggest problems. *Media and Communication*, 7(4):176–185, 2019.

[27] V. Corporation. Counter-strike: Global offensive, 2012.

[28] A. E. Depping and R. L. Mandryk. Why is this happening to me? how player attribution can broaden our understanding of player experience. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, page 1040–1052, New York, NY, USA, 2017. Association for Computing Machinery.

[29] B. Entertainment. Overwatch, 2015.

[30] C. Y. Foo and E. M. Koivisto. Defining grief play in mmorpgs: player and developer perceptions. In *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology*, pages 245–250, 2004.

[31] R. Games. League of legends, 2009.

[32] R. Games. Valorant, 2020.

[33] J. W. Hong, Q. Peng, and D. Williams. Are you ready for artificial mozart and skrillex? an experiment testing expectancy violation theory and ai music. *new media & society*, 23(7):1920–1935, 2021.

[34] A. Joardar. Examining changes in group acceptance of a newcomer from a different culture: An expectancy violation perspective. *International Journal of Cross Cultural Management*, 11(3):341–362, 2011.

[35] B. Kordyaka, S. Laato, K. Jahn, J. Hamari, and B. Niehaves. The cycle of toxicity: Exploring relationships between personality and player roles in toxic behavior in multiplayer online battle arena games. *Proceedings of the ACM on Human-Computer Interaction*, 7(CHI PLAY):611–641, 2023.

[36] Y. Kou. 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*, pages 81–92, 2020.

[37] Y. Kou and X. Gui. 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*, CHI PLAY '14, page 161–169, New York, NY, USA, 2014. Association for Computing Machinery.

[38] Y. Kou and X. Gui. Emotion regulation in esports gaming: A qualitative study of league of legends. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW2):1–25, 2020.

[39] S. Y. Lee. Expecting the unexpected? the moderating effects of intergroup expectations on being tolerated. 2021.

[40] C. M. Leigh Achterbosch and P. Vamplew. A taxonomy of griefer type by motivation in massively multiplayer online role-playing games. *Behaviour & Information Technology*, 36(8):846–860, 2017.

[41] M. Lemercier-Dugarin, L. Romo, C. Tijus, and O. Zerhouni. "who are the cyka blyat?" how empathy, impulsivity, and motivations to play predict aggressive behaviors in multiplayer online games. *Cyberpsychology, Behavior, and Social Networking*, 24(1):63–69, 2021.

[42] M. Märtens, S. Shen, A. Iosup, and F. Kuipers. Toxicity detection in multiplayer online games. In *2015 International Workshop on Network and Systems Support for Games (NetGames)*, pages 1–6. IEEE, 2015.

[43] T. MATTINEN. Toxic behavior in dota 2. 2018.

[44] C. McLaughlin and J. Vitak. Norm evolution and violation on facebook. *New media & society*, 14(2):299–315, 2012.

[45] A. Miller-Ott and L. Kelly. The presence of cell phones in romantic partner face-to-face interactions: An expectancy violation theory approach. *Southern Communication Journal*, 80(4):253–270, 2015.

[46] C. K. Monge and T. O'Brien. Effects of individual toxic behavior on team performance in league of legends. *Media Psychology*, 25(1):82–105, 2022.

[47] S. B. Nicholls and R. E. Rice. A dual-identity model of responses to deviance in online groups: Integrating social identity theory and expectancy violations theory. *Communication Theory*, 27(3):243–268, 2017.

[48] A. Ramirez Jr and Z. Wang. When online meets offline: An expectancy violations theory perspective on modality switching. *Journal of Communication*, 58(1):20–39, 2008.

[49] S. Sprecher. The role of expectations for liking and other positive affiliative outcomes in the get-acquainted process that occurs over computer-mediated video communication. *Current Psychology*, 42(2):1044–1057, 2023.

[50] K. P. Truong, D. A. Van Leeuwen, and F. M. De Jong. Speech-based recognition of self-reported and observed emotion in a dimensional space. *Speech communication*, 54(9):1049–1063, 2012.

[51] H. P. E. Valve Corporation. Dota 2, 2013.

[52] T. F. Waddell. A robot wrote this? how perceived machine authorship affects news credibility. *Digital journalism*, 6(2):236–255, 2018.

[53] K. D. Williams and B. Jarvis. Cyberball: A program for use in research on interpersonal ostracism and acceptance. *Behavior research methods*, 38:174–180, 2006.

Acknowledgment

Like many before me, I have faced many challenges—both academic and personal—during this research journey. I could not have done it without the help and support of many, many individuals. Though I cannot cover them all, I will try my best:

- First of all, I would like to thank my advisors, Professors Jeong-woo Jang and Juho Kim, for their guidance. I have learned so much from their great research insights and advice. Thank you very much for helping me grow as a researcher and seeing this research to fruition.
- I would like to thank my collaborators, Seoyoung and Yeon Su, for their work and feedback. Their contributions have made this research far more meaningful than I could have imagined by myself and their support helped me during the tough times.
- I would also like to thank my best friend Jeongeon for creating and sharing some of the best parts of my time at KAIST. I know she will do great no matter what she does.
- I want to express many thanks to the members of KIXLAB. Thank you, Haesoo, for being my amazing friend and role model. Thank you, DaEun, Seulgi, and Yoonsu for many fond memories. Thank you, Yoonseo, for always looking after me. I am also so grateful to Taesoo, Saelyne, Sangkyung, Kihoon, Hyoungwook, Hyunwoo, Jihyeong, Bekzat, and many others who have supported me and made each day brighter.
- I want to thank Joseph for many discussions and for showing me what good research can be. Thank you, Yoojin, for your kindness and feedback.
- I would like to thank Ms. Kim and other faculty and staff who handled many workload for my sake. Also, many thanks to all the participants who provided their insights for my research.

Finally, I give my deepest gratitude and love to my parents and sisters, who are my firmest believers. And, of course, thank you, Sej. You are my rock and the reason I keep going. I could have not done this without you.

Curriculum Vitae

Name : 이 주 훈

Date of Birth : February 16, 1998

Educations

2012. 8. – 2016. 6. Thomas Jefferson High School for Science and Technology

2016. 8. – 2022. 2. KAIST School of Computing (BS)

2022. 3. – 2024. 2. KAIST School of Computing (MS)

Career

2023. 3. – 2023. 12. KAIST School of Computing Teaching Assistant for CS101

Academic Activities

1. **Juhoon Lee**, Juho Kim, and Jeong-woo Jang. 2018. Look Upon Thyself: Understanding the Effect of Self-Reflection on Toxic Behavior in Online Gaming. In 2023 CHI ACM Conference on Human Factors in Computing Systems Workshop on Combating Toxicity, Harassment, and Abuse in Online Social Spaces.
2. Saelyne Yang, Sangkyung Kwak, **Juhoon Lee**, and Juho Kim. 2023. Beyond Instructions: A Taxonomy of Information Types in How-to Videos. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23).
3. Haesoo Kim, **Juhoon Lee**, Juho Kim, and Jeong-woo Jang. 2023. Towards an Experience-Centric Paradigm of Online Harassment: Responding to Calling out and Networked Harassment. In 2023 CHI ACM Conference on Human Factors in Computing Systems Workshop on Combating Toxicity, Harassment, and Abuse in Online Social Spaces.