Reducing Frustration in New and Novice Players in Competitive Games: An AI assistant for League of Legends Itemization

Master Research Thesis

Anesti Koliçi
a.kolici@students.uu.nl

April 2026

1st supervisor: Prof. dr. R.C. Veltkamp
2nd supervisor: Dr. A. Chatzimparmpas

Daily supervisor: Dr. J.D.L. Fijnheer

Department of Information and Computing Sciences
Faculty of Science

Universiteit Utrecht

Contents

Abstract

New and Novice players in competitive MOBAs like League of Legends often experience frustration due to the imbalance of flow, leading to low retention from players, who are unable to learn the complex mechanics. While existing Non-AI and AI tools provide item recommendations based on popularity, they lack the explanations, which can create player dependency. No studies go into detail of whether an AI assistant with explanations can specifically reduce frustration in new and novice players.

This research implements an AI 'Itemization' Assistant combining a Deep Q-Network model for item recommendations with a large language model for real-time explanations. Using a within-subjects design, 54 participants each played two 1v1 League of Legends matches, one with the AI assistant active and one without. Player experience was measured using the Game Experience Questionnaire and system usability assessed with the System Usability Scale questionnaire.

The assistant significantly reduced frustration and improved overall player experience. All four GEQ sub-scales showed significant improvements in the AI condition. The results suggest that frustration reduction is attributed to the explained by the AI Assistant, independent of practice effects.

Overall system usability was the strongest predictor of frustration reduction, and participants showed a clear preference to explanations, rather than recommendations. Knowledge gain was significant overall, but improvement was concentrated among beginner and intermediate participants, while the majority of those who had never played before, showed no change.

1 Introduction

In competitive MOBA games, like League of Legends, item selection plays a critical role in performance. Understanding when and why to choose specific items requires game knowledge and experience. Players may struggle to make effective decisions and become familiar with the mechanics, affecting both enjoyment and success. When it comes to being a new player in competitive games, the learning curve can be steep. These players often encounter hard mechanics, complex itemization systems, and high-pressure environments filled with toxicity, which can create frustration and eventually lead to quitting the game.

To address this issue, an increasing number of researchers are exploring the potential of intelligent systems that function as virtual trainers or in-game assistants. These systems not only aim to support decision making, but also to reduce player frustration and increase long-term engagement. However, most existing solutions focus on high-level analytics or post-game feedback. Real-time options are being introduced into different systems. In this case with a focus on new and novice players.

This research proposes the design and evaluation of an AI-powered in-game assistant that provides live itemization recommendations based on game data. The assistant also offers understandable explanations for its suggestions, with the aim of supporting learning and reducing frustration. Unlike existing training systems, this assistant is intended to function during gameplay, offering advice and suggestions based on the needs of players with limited game knowledge.

The study focuses on a simplified version/matchup of the game: a 1v1 custom mode on the top lane, limited to a single champion belonging to the same role. This reduces the variables and isolate the impact of item suggestions.

When it comes to training data for the assistant model, it's collected using the Riot Games API, including detailed statistics such as champion matchups, item builds, win rates, gold economy, etc. The assistant is able to recommend items based on the in-game situation. An LLM is then used to generate explanations for each suggestion made.

2 Background and Context

In order to understand what can lead players to frustration while playing games, this research explores the psychological frameworks of flow and frustration. Flow state is balanced by challenge and skill, if there is no challenge it can lead to boredom, but on the other hand a new player that faces too great of a challenge from the start can experience frustration. The competitiveness of League of Legends, and competitive games as a whole, is a factor to consider that can increase frustration. This section will first discuss the Flow Theory and its relationship to games, how higher challenge without improvement can lead to high frustration, and quitting of games. Choosing League of Legends as the game to focus on, arrived from that extensive research that the game has been used on, as well as the tools that allow for the gathering of data. The research focuses on Itemization, as it allows for a better controlled environment for the experiment and is one of the core mechanics that new players have a hard time with. The AI assistant contains both recommendations and explanations, since a gap was identified from other non-AI tools widely used. Other research done on similar concepts of AI assistants, don't fully explore the possibility of having it as an assistant to learn the game, or the least provide only tips.

2.1 Flow Theory

The 'Flow Theory', introduced by Mihály Csíkszentmihályi, refers to "a state of concentration so focused that it amounts to absolute absorption in an activity". In this state, individuals experience a sense of effortlessness, control and timelessness. Several factors must align in order to reach a state of flow that includes a goal, immediate feedback, and a balance between challenge and skill, bringing the "optimal experience". As such, the flow state plays an important role in the enjoyment and long-term engagement of the game.[1]

In the context of video games, flow theory has become an important part in creating a great player experience, with a lot of research around it. Games are a medium that can facilitate flow states through their mechanics, immediate feedback and a goal for players to focus on. Researchers have developed approaches for analyzing these game elements to support or hinder the optimal experience.[2]. Some games aim for a carefully crafted experience, challenge and progression, and thus creating a "Flow state Zone".[3]. A model to mention would be 'GameFlow', it identifies eight key elements for evaluating player enjoyment and flow in games: concentration, challenge, skills, control, clear goals, feedback, immersion, and social interaction.[4]. Arguably, the two most important elements are challenge and skill, a balance is needed in order to get into the flow zone. Having the challenge be higher than the skill can lead to anxiety, frustration and in some cases giving up, while having the challenge be too easy can lead to a boring experience to a good player.[1] As such, regulating flow state is important for long-term enjoyment of games.

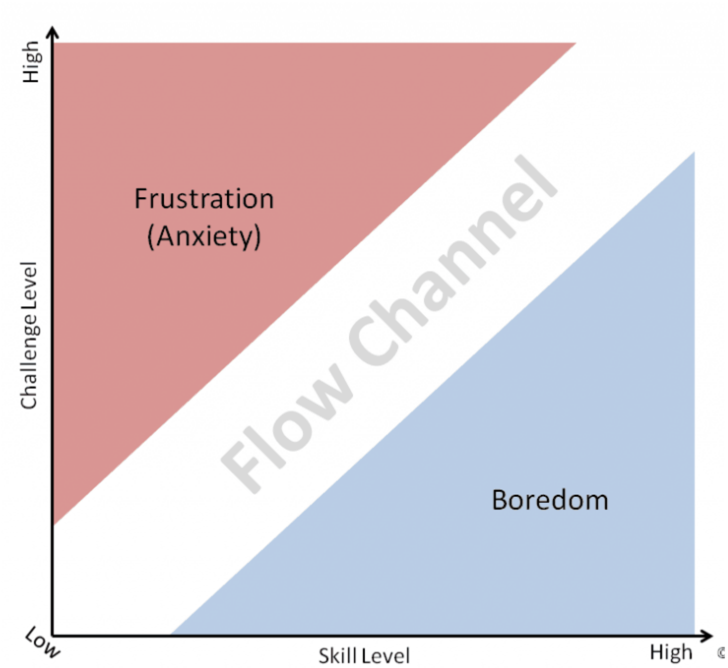

Figure 1. *Simplified Flow Theory Graph*

When it comes to competitive games, maintaining a balanced flow can be very challenging, especially in the Multiplayer Online Battle Arena(MOBA) genre. Research made on games across different genres has found that MOBAs produce less immersion while simultaneously creating stacking challenge and frustration levels. This genre is for players who want a sense of satisfaction from teamwork, competition and mastery of complex gameplay mechanics. This presents a paradox where the very complexity that makes MOBAs engaging for experienced players becomes a barrier that hinders and stops the new player engagement. Complex decision-making in games such as League of Legends, including itemization, champion mechanics, and strategic positioning, can overwhelm new players and push the challenge beyond their current skill level, leading to negative responses, such as frustration. Studies measuring flow experiences in League of Legends have shown how challenging it can be to achieve consistent flow rates. These difficulties suggest that flow in MOBAs is more complex and fragmented compared to other genres.[5]. Particularly with League of Legends, research has shown that the flow states and experiences may depend on factors outside of the gameplay, such as out of game immersion, matchmaking, microtransactions etc.

The real-time nature of League of Legends, makes it even harder to adjust to new information, losing flow. Unlike singleplayer games, where players can take their time playing and processing the information, League of Legends matches require focus and immediate responses to changing situations. The toxicity often displayed in competitive games can make it even harder for someone to achieve optimal experience, raising frustration. When a new player starts League of Legends, they typically find themselves in the anxiety and 'Frustration Zone' of the flow model. Especially at the start where their skill can be considered zero and the challenges, such as characters and itemization, at the peak. As players develop their skills through practice and experience, it is expected to get closer and closer to the flow state zone, eventually reaching it. When players can't achieve or maintain the flow state, negative emotional responses develop. The most prominent one in competitive games is frustration, which can lead to anger, sadness, etc.

2.2 Frustration

Frustration is one of the key challenges a new player must overcome in League of Legends and understanding the emotional response is essential to design a system that prevents it. Frustration, in its common context, is understood as a negative emotional response coming from the inability to achieve a goal or satisfaction from the effect of obstacles or problems. Traditionally, frustration theory suggests that this emotional state is an outcome of blocking goal-driven behavior, leading to disappointment, irritation and helplessness.[6] However, when it comes to frustration in the gaming context, it differs in some ways. Instead of frustration being a passive response to blocked goals, in gaming it becomes active and dynamic. Players interacting with the game, the design of it and the flow of the challenge, all can be factors that affect frustration.[7]

Research has shown that frustration and flow are in opposition to each other, and has proven useful in understanding player experience. The optimal experience occurs when there is a balance of challenge and skill, when the challenge significantly exceeds the player's skill level, it leads to unwanted negative emotions, represented by frustration. This opposition can be seen more in games that require real-time decisions and immediate feedback, particularly MOBAs. In these games, the balance might shift while playing, going between states depending on their skill and demands of the game.[8]

MOBAs have particularly high frustration levels among players, which can be attributed largely to the competitive nature and the difficulty of the different mechanics. Some sources of frustration shared among MOBAs are: complexity overload, knowledge gaps, lack of feedback to improve on, skill mismatch, including Itemization. Research has also demonstrated that these sources create a cyclical pattern, where frustration can snowball from a small inconvenience. This creates a growing cycle where initial frustration from game mechanics leads to negative responses, which in turn generate more frustration and ruin the experience all together.[9]

2.3 Retention

Player retention is one of the problems MOBAs are facing, partly because of frustration. Academic research done on League of Legends using survival analysis, revealed that most of the new players do not continue beyond the initial learning period with the first weeks being the most important for player retention. The study also examined the willingness for players to continue playing MOBAs and several factors were observed, frustration being the primary predictor for quitting the game.[10].

When a new player encounters complex mechanics like Itemization, champion abilities and macro mechanics without any initial guidance or help, can create an imbalance between challenge and skill. This imbalance pushes players into the frustration zone of the flow model, which led to negative experiences that get worse with time. As mentioned, frustration can directly decrease retention, but also influences other factors that could keep the player engaged. By addressing one of the mechanics, in this case Itemization, overall frustration levels could be reduced during the critical period and potentially improve retention by making the initial learning process less overwhelming.

2.4 League of Legends

League of Legends is a popular Multiplayer Online Battle Arena (MOBA) game developed by Riot Games and released in 2009. In the game, two teams of five players compete to destroy the opposing team's 'Nexus' (Base), also known as the base, while taking down Turrets (Towers) on the way. The matches are played on a symmetrical map called 'Summoner's Rift', divided into three lanes ('Top', 'Middle', and 'Bottom'), along with the 'Jungle', an area filled with neutral objectives. Each player selects a unique champion from a roster of over 160 characters, with each champion offering distinct abilities, playstyles, and strategic roles such as tank, support, mage, marksman, etc.

Gameplay in League of Legends focuses on real-time decision-making, map awareness, coordination, and adaptability. Players earn gold and experience by defeating enemy champions, minions, and jungle monsters, allowing them to level up and purchase items that make them stronger. The itemization system is deep and situational, requiring players to make informed choices based on their champion's strengths and their opponents' weaknesses.[11]

Selecting League of Legends as the game for this research is based on several factors that distinguish it from other MOBAs. Research shows that League of Legends is the most studied MOBA in academics, providing more specific existing research to work with.[12] Another thing to mention is the player base, 140 million monthly users, compared to the second most played MOBA at 11 million. Considering the numbers, there are still a lot of people that have never played or have barely done so, giving good data collection opportunities.

League of Legends has some clear advantages when it comes to accessibility as it is less complex. Although still hard and complex, reviewers mention that it's easier to pick compared to other MOBAs, making it more beginner-friendly. This accessibility is crucial to consider with the flow states, as something overly complex can push a new player so much into the anxiety and frustration zone that even removing one of the mechanics might not decrease the learning curve by much.

2.5 Itemization

'Itemization' in League of Legends refers to the process of buying and upgrading equipment, referred to as 'Items'. They serve to enhance several attributes of a champion, including abilities, stats and passive effects. There are over 200 unique items in the game that require Gold, earned through slaying PvE NPCs, other players' champions and destroying structures. The items in the game work in a tiered system, where normal items can be combined to make up to legendary items. This requires players to plan their build path in advance depending on the match. What makes Itemization complex is the factors that affect it, such as the role of the player, team composition, gold available, PvE NPCs and many more. For example, a player might prioritize armor if the enemy team leans more into physical damage, while another player will buy items with more damage to boost their stats more and try to get a larger advantage.[11]

The challenges created by the number of options and complexity can overwhelm a player. The game requires understanding about item synergies, team compositions and build paths. Quick decisions are important, as every second spent in base is wasted, increasing the pressure all while having to manage other complex mechanics in the game, making an AI assistant very useful.

2.6 Related Works, AI Tools in Gaming

Artificial Intelligence has transformed gaming, evolving from simple scripts and systems that adapt to every player's action, to providing personalized tips in some cases. Especially in the competitive scene, AI is being used in analyzing games and matches, predicting outcomes and personalizing or enhancing player experience. In strategy games, mainly RTS and MOBA, the development of such systems is getting more common. While older iterations of AI in games focused on creating "real" NPCs, current tools have shifted their focus toward player assistance and learning enhancement.[13]

As mentioned, early implementations include hint systems and tutorials, but most modern assistants use machine learning to provide suggestions in real time. The Nemesis System in the Shadow of Mordor showed how AI can create dynamic and responsive experiences, while other games like Forza implemented AI racing lines that adapt to the player's skill. It is very important to consider what all of these solutions have in common, they do not disturb the flow state of the game, by making sure it's seamless and the assistance is provided in a non-intrusive manner through visual overlay, audio, etc. However, most successful implementations are from single-player games, this can be attributed to the pacing and the non-competitive nature of these games.[14 15]

When it comes to analytics, the space is quite diverse in what it offers. Platforms like Mobalytics, U.GG and OP.GG are the most used when it comes to player improvement. U.GG and OP.GG primarily focus on statistical aggregation and visualizing said statistics. They also offer Item recommendations, but they are based on win rates and popularity across all skill levels without taking into consideration the player's skill level or the situation the player is in. These recommendations tell you what to buy, but not when and why, giving no context to the player to learn from. Mobalytics is essentially the same, but it is able to provide all the information after a game has ended, as well as a simple overlay with item recommendation. This approach also has its limitations, players receive feedback after the game has ended and the learning opportunity is no longer there. It can be said that a player can study this data themselves, but in most cases this takes as long as the game itself, as it needs a full game review.

What all of these platforms have in common is that they lack the explanations needed for a new player to learn and understand the game. Another thing to add is that they create a dependency on them, they provide answers without explanations so the player doesn't have to waste time. Although for experienced players, an immediate decision is the difference between winning and losing a game, a new player cannot improve long-term as fast. This phenomenon is also referred to as "guided incompetence", players will win more games at the cost of being unable to make informed decisions independently.[16]

AI assistance, specifically, in gaming has been limited, but shows some promising results. Research done on the paper A Tutor Agent for MOBA Games demonstrated that players using their assistant showed performance improvements compared to the rest. Their system relied on a support character that provided both gameplay assistance and tips on how to improve. Players showed a lot of satisfaction, but it is important to mention that the experiment was limited by their participant number, which was only 6. Their approach provided general gameplay tips and guidance, but didn't address the specific gaps in players' knowledge, as well as didn't provide explanations regarding the tips.[17]

Similar AI systems researched on RTS(Real-Time Strategy) games have also been researched, particularly RTSMATE. RTSMATE gives real time advice during gameplay and it was developed using decision trees made by expert guides. The system could provide tactical and strategy tips, while also having features to predict the technology progress of the opponent. Similarly, the experiments with the players showed significant increase in performance, but could be argued that some dependency was created from it. It also didn't feature any explanations or context for the tips, as such the paper didn't address the cognitive challenges that new players have to overcome from learning complex game mechanics.[18]

Research involving LLMs and recommendations suggests that people also prefer recommendations done in this way. In this case Llama 2 was used to generate movie explanations and compared them to the traditional template method.[19] The study measured effectiveness, persuasiveness, transparency, trust, and overall satisfaction, which all showed increases compared to the traditional method. Majority of the participants(82.5%) preferred the generated explanations, and having them being understandable is also a core part of them. Similarly, another study comparing generic user-based explanations to ones generated by ChatGPT 3.5.[20] While they also found similar results to the measured dimensions, the user-based explanations showed that they weren't perceived as more personalized or effective. This analysis revealed that explanation effectiveness is primarily predicted by persuasiveness and personalization, with persuasiveness being the stronger predictor.

Currently, AI assistance tools focus on performance and gameplay optimization, but pay little attention to reducing player frustration. While existing research shows that gaming frustration affects player retention, and AI assistance tools can improve performance, no studies have examined whether it specifically reduces frustration in new players. This represents a gap, because performance improvement may not translate into reducing frustration. Furthermore, current tools provide recommendations without explanations, leaving players reliant on them rather than helping and reducing frustration.

The focus of the ai assistant itself is to improve the experience and reduce frustration by improving the learning curve, one less mechanic for the player to get frustrated with. Research on other systems, such as Intelligent Tutoring Systems, has shown that explanatory feedback can significantly improve learning outcomes compared to simple correct or incorrect feedback. In a particular research, Korbit, a large-scale dialogue-based ITS, demonstrated that personalized hints and explanations, as well as Wikipedia-based explanations, help improve student learning outcomes significantly.[21]However, unlike traditional ITS, which operate in low-pressure educational environments, this AI assistant needs to function in a "high-stress" competitive environment. The learning aspect of the assistant distinguishes it from other game performance and optimization tools, putting it more in line with learning assistance designed to improve the overall experience and reduce frustration, in hopes that new player retention improves.

3 Research Questions

The research questions emerged from the identified gaps in current assistance tools for competitive gaming, as well as frustration measurement approaches in MOBAs. As established in the literature review, current tools provide item recommendations but lack any explanatory capability. Flow theory suggests that reducing the skill-to-challenge gap should also decrease frustration, but no studies have specifically tested if an AI Itemization Assistant with explanations can achieve this outcome. The questions address these knowledge gaps by examining both the AI assistance effects on player frustration and player experience.

Research Goal: To gain insights into what an AI Assistant reveals about improving player experience and retention in new and novice League of Legends players.

Research Question: Can an AI Itemization Assistant with explanations, for item recommendations in League of Legends, reduce frustration levels and improve player experience in new and novice players.

- **Sub-question 1:** What design elements and user interaction patterns of the AI assistant are improving the player experience and reduce frustration?

  This question aims to find the optimal design principles for the AI assistance. Since the assistant is real time, it was necessary to understand which UI design, visual layout, method of information presentation, etc. is best for new players. As such, learning the effectiveness of the assistant and provide a design guideline for future tool similar to this.

- **Sub-question 2:** Do players using the AI assistant improve retention levels and does it relate to frustration and experience?

  This question directly investigates if the AI assistant improves retention intentions in new and novice MOBA players. As mentioned in the literature review, frustration is an important factor for predicting quitting in League of Legends, particularly as a new player. By comparing the frustration levels to retention intentions, evidence can be gathered that shows whether the assistant was successful and had an impact on the experience of the players.

- **Sub-question 3:** What aspects of the AI assistant do players perceive as relevant for improving and reducing their frustration?

  User perception can be varying and understanding it is essential in developing effective tools in game environments. This question aims to find if the recommendations, explanations, the UI, and other elements of the AI assistant are the most effective in reducing frustration.

- **Sub-question 4:** Can the AI assistant with explanations influence the players' understanding of the Itemization mechanic?

  Explanations are an important part of this AI assistant, which is what makes it different from publicly released tools. By studying the effect of the assistant with knowledge gain, the research aims to identify a better solution for learning complex games, such as MOBAs.

4 Methodology

The methodology of this research is designed to evaluate the effectiveness of the AI Itemization assistant for new League of Legends players. The research uses a within-subjects experimental design, collecting quantitative data to assess the assistant's impact on player frustration and experience. Each participant plays two controlled 1v1 League of Legends matches: one with the AI assistant active (treatment) and one without it (control). The order of conditions is divided into two groups, participants that play the control match first and those that play the treatment match.

When it comes to developing the AI assistant, it is divided into a Deep Q-Network (DQN) model responsible for generating item recommendations, and a large language model (LLM) that produces natural language explanations for each recommendation. Data for training the DQN model is collected through the Riot Games API, while the LLM component uses Claude accessed via the Anthropic API to generate real-time explanations. An overview of the system architecture is shown in Figure 2.

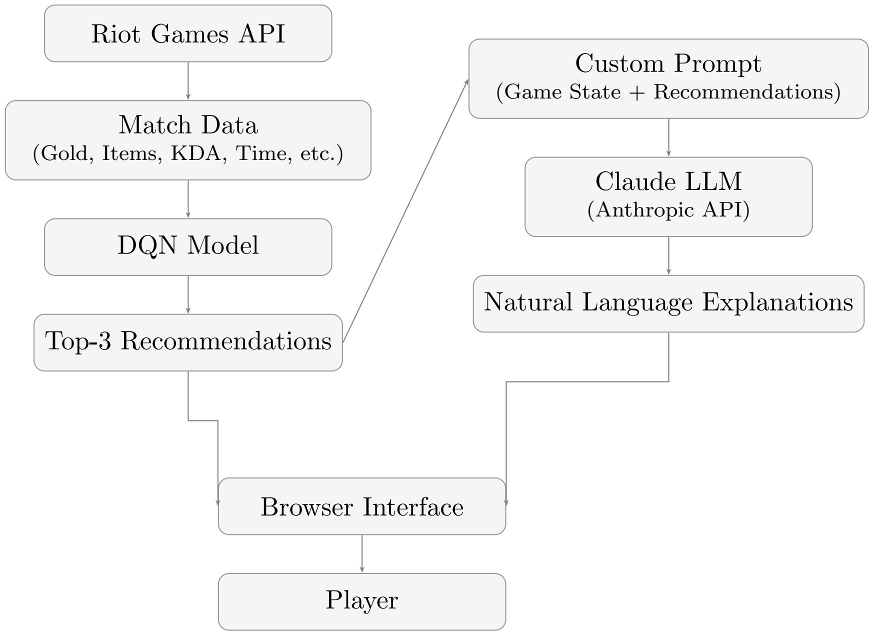

Figure 2. *System Architecture*

4.1 Model Development

The AI assistant uses a two-component architecture that separates the recommendations from the explanations, creating two distinct systems. The DQN model handles item recommendations, while the LLM then explains those suggestions in natural language. Users are able to see the recommendations in a separate browser window. This separation allows each system to be developed, validated, and updated independently.

4.1.1 RIOT API

Training data for the DQN model was collected using the Riot Games API, which provides access to detailed match and player statistics. The dataset consists of tens of thousands of matches from the Platinum to Diamond skill tiers, specifically filtering for matches where the champion Sion is played in the Top lane. This scope was chosen to keep the training distribution consistent with the experimental setup while still providing a large and representative sample of high-quality gameplay.

Key parameters extracted per match include:

- Gold earned and gold spent at specific game timestamps
- Complete Item purchase history including timestamps

- Player Stats(KDA, cs, damage done, etc.)
- Final match outcome (win or loss)
- Champion matchup data for the Top lane(opponents)

Live game state data is fetched during the gameplay through League of Legends Client API. It uses a port (2999) to get the real-time information about the ongoing match instantly. Item metadata, descriptions, costs and synergies, are sourced from CommunityDragon, an open repository of League of Legends game assets.

4.1.2 DQN MODEL

For the Item recommendation system, a Deep Q-Network (DQN) architecture was selected. A Deep Q-Network is a reinforcement learning architecture that combines Q-learning with a deep neural network.Q-learning is a model-free algorithm that measures the quality, in this case known as Q-value, of taking an action in a given state. This value represents the expected reward an agent receives. In standard Q-learning the values are stored in a lookup table, while DQN replaces this table with a neural network, enabling generalization across large state spaces.[22]

DQN is well-suited for several reasons: Item selection in League of Legends takes up a discrete action space, as a player must choose from a set of items at each decision point, which fits perfectly with the discrete action formulation that DQN is designed for. The model learns Q-values for each possible item choice given the current game state, enabling it to rank the items and output a list of recommendations. The model is able to output more recommendations but for the purposes of this study the top 3 are selected, as shown on Figure 3.

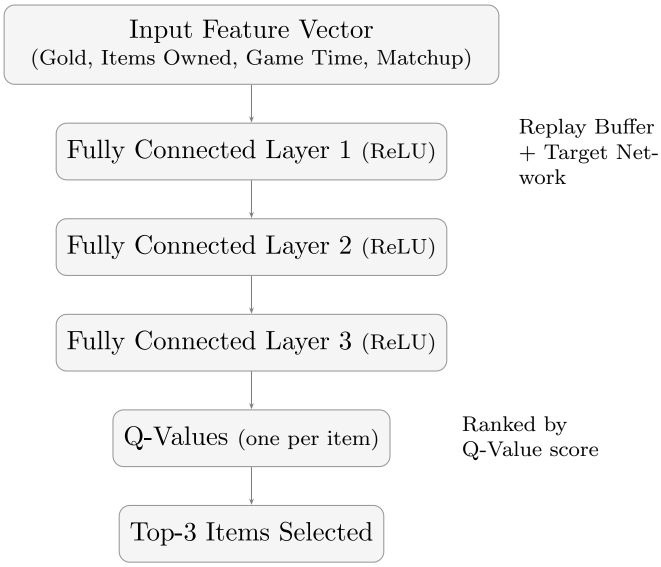

Figure 3. *DQN Model Architecture*

The model relies on offline learning from historical match data(RIOT API). The states of the matches fed to the model get encoded to a feature vector, including current gold, previously purchased items, elapsed game time, and champion matchup context. The model was trained exclusively on Sion Top lane data as he was the character that participants played, as such resulting in more focused and accurate predictions.

The model was trained over 100 epochs. Figure 4 shows the training loss and Top-3 validation accuracy curves across training. The loss curve shows a sharp spike at epochs 1–2, which is a characteristic of DQN training. From approximately epoch 40, the loss stabilizes at around 0.10, indicating that the network has reached a consistent and stable Q-value estimate, no longer making large changes.

The Top-3 validation accuracy follows the expected pattern of high early volatility—swinging between 0% and 70% in the first 15 epochs as the replay buffer fills and the exploration rate decays—before converging to a stable range. The model achieved a best Top-3 accuracy of 81.3%, with accuracy consistently holding above 70% in the final training epochs. Top-3 accuracy is the appropriate evaluation metric here because the assistant shows participants three item recommendations rather than a single prediction. When the correct item appears among the three suggestions, the recommendation is considered successful from a practical standpoint.

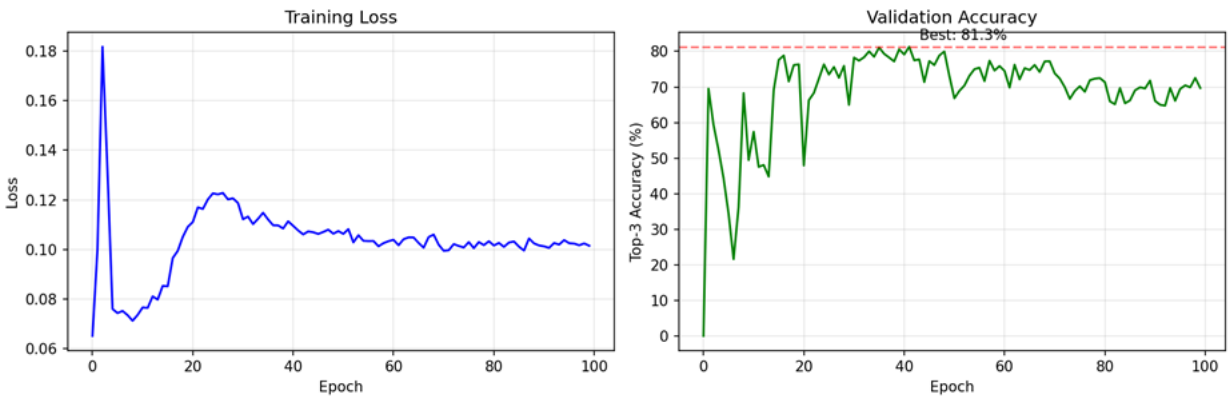

Figure 4. *Training Loss and Top-3 Accuracy*

4.1.3 LLM Explanation System

To generate the explanations for each item recommendation, the system integrates Claude via the Anthropic API. The LLM receives a custom prompt containing the relevant information, which include the recommendation given by the DQN model. It then produces an explanation for each suggested item, describing why the item is appropriate for the current game.

Claude was selected over locally hosted alternatives, including DeepSeek, GPT and Gemma models, following a testing phase of response quality and latency. Running a capable language model locally introduced too many delays for a real time gameplay experience, whereas the Anthropic API delivered significantly faster response times. Response quality was also evaluated to be higher with Claude, particularly in terms of explanation clarity and context.

Explanations are displayed to participants via a browser window interface that runs alongside the game, as shown on Figure 5. The interface shows the three recommended items, their confidence scores, and a plain-language rationale for each. This design ensures the assistant is informative without requiring participants to leave the game client or interrupt their decision-making process excessively.

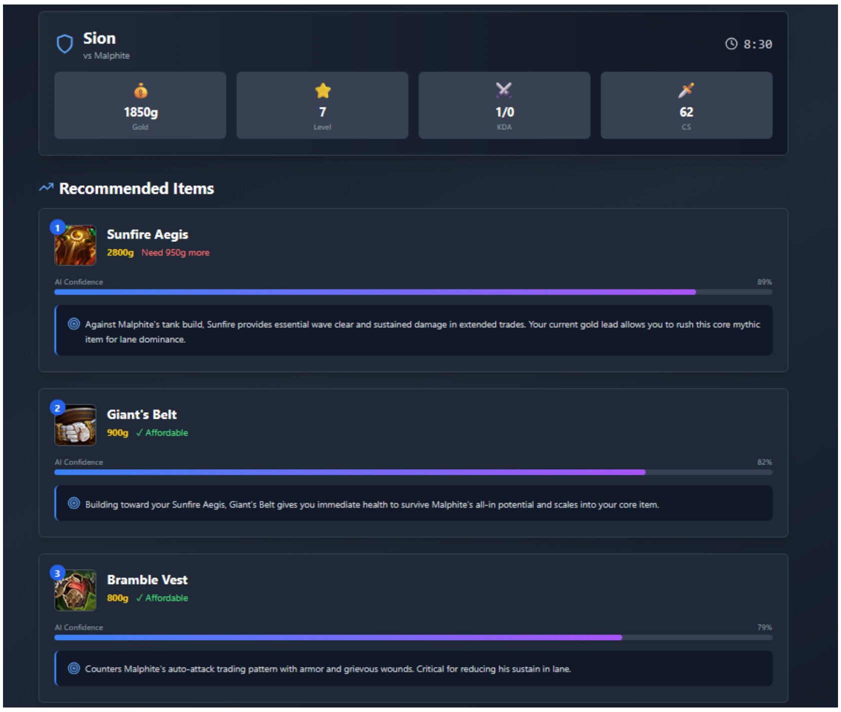

Figure 5. *Interface*

4.2 Experimental Setup

4.2.1 Experiment Design

To obtain the necessary data on the effect of the AI assistant, an experiment was conducted where players played two matches of League of Legends. The experiment uses a within-subjects design in which every participant plays under both conditions: one match with the AI assistant (treatment) and one match without it (control).

Condition order is balanced, meaning about half of participants play the control match first, followed by the treatment match, while the other half play in the opposite order. Survey questions are asked between the two matches to get the data on frustration and experience measurements that are specific to the first match before the participant moves on to the second.

Statistical power analysis

Assuming the Effect Size (Cohen's d) is on a medium level of intervention, it was set at d=0.5. Alpha and Power are standard at 0.05 and 0.8(80% of detecting an effect). Calculation was done using the formula for a paired t-test, giving the a target minimum of 32 participants. During the experiment a uniform number 40 participants was targeted, with more participants invited when possible.

$$n = \frac{(z_{\alpha/2} + z_\beta)^2}{d^2} = \frac{(1.96 + 0.84)^2}{0.5^2} = \frac{7.84}{0.25} = 31.36 \approx 32 \tag{1}$$

A pilot study was conducted with five participants prior to the main experiment to identify any issues, refine survey wording, and verify that the technical setup functioned properly.

Participants were recruited from the student population at Utrecht University and campus ads, student housing and online communities. A small reward was given at the end to motivate participation. They had to have less than 10 hours of total playtime in League of Legends and be over 18 years old.

Experiment Implementation

Each experiment session was conducted on site or online and lasted approximately 60 minutes in total. All matches were played on a custom 1v1 match on Summoner's Rift, Top lane only, against an intermediate difficulty bot opponent that stays the same. The champion pool is restricted to Sion, a fighter-class champion,to keep the recommendation model focused. Participants are not required to have prior experience with playing Sion. When it comes to choosing the items, they were instructed to use the in game shop, unless playing with the Assistant, which was required by them to use it to make their choices. Figure 6 presents a simple timeline of the experiment.

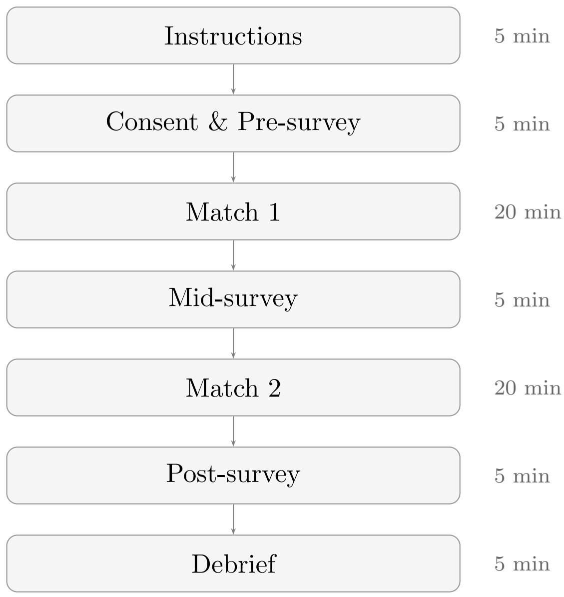

Figure 6. *Experiment timeline. Total estimated duration: 60 minutes.*

Several Variables have to be kept in mind for the experiment. Table 1 shows the identified Independent, Dependent and Control Variables.

Table 1. *Experimental Variables*

Type	Variable	Measurement
Independent	AI assistant condition	Present vs. absent
Dependent	Player frustration	GEQ Frustration subscale (pre, mid, post)
	Player experience	GEQ subscales: Competence, Flow, Challenge (post each match)
	Itemization knowledge	Custom 5-item MCQ (pre and post session)
	System usability	System Usability Scale (treatment group only)
	In-game performance	KDA ratio, gold earned, items purchased (automatic)
	Retention	Likelihood to continue playing (post-session survey)
Control	Champion	Restricted to Sion for all participants
	Match format	1v1 vs. intermediate bot, Top lane only
	Match duration	Capped at 20 minutes
	Game settings	Standardised across all experiment machines
	Condition order	Counterbalanced using predetermined scheme

5 Survey

In order to assess the frustration and experience of the player, the Game Experience Questionnaire (GEQ) was used. It was developed by IJsselsteijn, de Kort, and Poels and is a questionnaire used widely in the field of game research.[23 24] A version of it that mainly measures frustration and experience and was used both on the pre- and post-game surveys, across a 5-point Likert scale. The post-game survey also included measurements for Competence, Flow and Challenge to further evaluate the overall player experience. Understanding related to the Itemization system was measured through a custom knowledge survey both before and after the playing, consisting of five multiple choice questions regarding the system. This pre-post design allowed measurement of knowledge gains and directly address whether the AI assistant's explanations help learning as well.

When it comes to the match played with the AI Assistant, some additional questions were asked regarding the effectiveness of the AI. The System Usability Scale (SUS) is a validated questionnaire by Brooke, a modified version of it was used to assess the overall usability of the AI assistant.[25 26] Custom questions were included to measure quality, trust in the recommendations and perceived value. Finally, participants completed some short questions regarding retention, their likelihood to continue playing MOBAs and whether their interest regarding the genre increased, thus providing the measurements to assess the assistant's impact on player engagement and retention.

6 Results

6.1 Preliminary Analyses

A total of 54 participants completed both experimental sessions and were included in the final analysis. Table 2 presents the characteristics of the sample.

Table 2. *Participant Characteristics ($N = 54$)*

Variable	Category	n	%
Age			
	18–24	30	56
	25–34	21	39
	35–44	3	6
Gender			
	Male	30	56
	Female	24	44
Occupation			
	Full-time student	32	60
	Employed	19	35
	Unemployed	3	6
League of Legends experience			
	Never played	28	52
	Beginner	15	28
	Intermediate	11	20
Gaming hours per week			
	0–1 hours	18	33
	2–5 hours	17	31
	6–14 hours	11	20
	15+ hours	8	15
MOBA experience			
	No prior MOBA experience	22	41
	Some prior MOBA experience	32	60
Preferred gaming platform			
	PC	16	30
	Smartphone	17	32
	Console	15	28
	Mobile console	4	7
	Other	2	4

The largest age group was 18–24 (n = 30, 56%), with the majority being male (56% male, 44% female). Most were full-time students (60%), followed by employed professionals. Over half reported never having played League of Legends (52%), with the rest split between beginner (28%) and intermediate (20%) players. A majority reported some prior MOBA experience (60%). Gaming hours were distributed across the full range, with approximately two-thirds of participants playing five hours per week or fewer. Platform preference was spread across PC (30%), smartphone (32%), and console (28%).

The internal consistency of all scales was measured with Cronbach's alpha. As shown in Table 3, all scales exceeded the minimum acceptable threshold of $\alpha = .70$.[27] Normality of GEQ subscale change scores was measured using the Shapiro-Wilk test. Where normality violated ($p < .05$), non-parametric Wilcoxon signed-rank tests were used for the within-subjects comparisons in place of paired t-tests.

Table 3. *Cronbach's Alpha Reliability Coefficients for All Measurement Scales*

Scale	**Cronbach's α**	**N Items**
GEQ Negative Affect	.76	8
GEQ Competence	.96	10
GEQ Immersion	.91	8
GEQ Challenge	.90	10
SUS (reversed items)	.93	10
AI Explanations	.88	4
AI Recommendations	.83	3
Timing & Disruption	.77	3
Perceived Impact	.85	3

Note. All scales exceeded the minimum threshold of $\alpha = .70$

6.2 Main Research Question: Frustration and Player Experience

Can an AI Itemization Assistant with explanations, for item recommendations in League of Legends, reduce frustration levels and improve player experience in new and novice players.

To evaluate whether the AI assistant reduced frustration and improved the overall player experience, GEQ subscale scores were compared between Match without AI and Match with AI using Wilcoxon signed-rank tests. Descriptive statistics are presented in Table 4 and test results in Table 5.

Table 4. *GEQ Subscale Descriptive Statistics by Match Condition*

	Match without AI		**Match with AI**	
Subscale	**M**	**SD**	**M**	**SD**
Negative Affect	2.4	0.8	2.0	0.7
Competence	2.4	0.9	3.1	1.0
Immersion	2.5	0.8	3.0	1.0
Challenge	3.0	0.8	2.7	0.6

Note. GEQ scale range: 1–5.

Table 5. *Wilcoxon Signed-Rank Test Results for GEQ Subscales ($N = 54$)*

Subscale	**Z**	**p**	**r**
Negative Affect	−2.7	.008	.4
Competence	−5.8	<.001	.8
Immersion	−3.9	<.001	.5
Challenge	−5.0	<.001	.7

All four GEQ subscales showed statistically significant differences between the two match conditions. Negative Affect being the with the lowest effect, falling at the medium range, while Competence, Immersion and Challenge had significant improvements when comparing the matches without AI and with AI respectively.

To assess whether the two matches played in different orders influenced results, an independent-samples t-test compared GEQ scores between the two groups. Table 6

Table 6. *Order Effect: Independent Samples t-Test on GEQ Change Scores by Condition Order ($N = 54$)*

Variable	**M (SD)**		**t**	**df**	**p**
	Without AI first	**With AI first**			
Δ Neg. Affect	−.8 (1)	+.1 (.7)	−3.9	48	<.001
Δ Competence	+.7 (.5)	+.8 (.7)	−.3	43	ns
Δ Immersion	+.4 (.9)	+.6 (.6)	−.7	47	ns
Δ Challenge	−.4 (.5)	−.4 (.4)	−.1	52	ns

Participants who played without AI first showed a mean frustration decrease of −.8 points in the AI match ($M = .8$, $SD = 1.0$), whereas participants who played with AI first showed a slight frustration increase when playing the next non-AI match ($M = +.1$, $SD = .7$). There was no statistical significance in the order effects that were observed for competence, immersion, or challenge (all $p < .05$, indicating that the improvements on these sub-scales were unrelated to condition order.

6.3 Sub-Question 1: Design Elements and Frustration Reduction

What design elements and user interaction patterns of the AI assistant are improving the player experience and reduce frustration?

Table 7 presents the Pearson correlations between each design element and frustration change (Δ Negative Affect), which helped with understanding the importance of each system.

Table 7. *Pearson Correlations Between Design Elements and Frustration Change (Δ Negative Affect)*

Design Element	**r**	**p**
System Usability		
SUS Score (overall usability)	$-.7$	<.001
Recommendations		
Rec1 — Recommendations helpful	$-.41$	.002
Rec_Mean — Recommendations mean	$-.40$	.003
Rec2 — Trust recommendations	$-.34$	.012
Rec3 — Better decisions	$-.28$	.038
Explanations		
Exp1 — Easy to understand	$-.34$	.013
Exp_Mean — Explanations mean	$-.33$	.014
Exp2 — Helped me learn	$-.30$	.028
Exp3 — Clear and well-written	$-.29$	.036
Exp4 — Applied to decisions	$-.23$	ns
Timing & Disruption		
Tim_Mean — Timing mean	$-.25$	ns
Tim1 — Appropriate timing	$-.19$	ns
Tim2 — No disruption	$-.19$	ns
Tim3 — Easy during game	$-.23$	ns

Note. Negative r indicates higher design ratings are associated with greater frustration reduction.

Recommendation ratings showed the most meaningful results with frustration change, with Rec1 (helpfulness of recommendations) reaching $r = -.41$ ($p < .05$) and Rec_Mean reaching $r = -.40$ ($p < .05$). Explanation ratings also helped significantly with frustration reduction, with Exp1 (ease of understanding) at $r = -.34$ ($p < .05$) and Exp_Mean at $r = -.33$ ($p < .05$).

The mean SUS score is $M = 68$ ($SD = 18.5$), which falls exactly at the average benchmark, indicating possible improvements.

6.4 Sub-Question 2: AI Assistance and Player Experience and Retention

Do players using the AI assistant improve retention levels and does it relate to frustration and experience?

Retention intentions, measured through replay likelihood averaged M = 2.7 (SD = 1.4) and recommendation to a friend averaged M = 2.8 (SD = 1.2), both below the midpoint of 3, suggesting modest overall retention intentions. It is worth noting that although the retention didn't change much in most the players, there was observed that those who experienced greater negative affect reduction also reported a greater intention to play again.

Interest in MOBA games remained largely stable following the study. Most of the participants (56%, n = 30) reported no change in interest, while 20% reported a decrease and 17% reported an increase. The full distribution is presented in Table 8.

Table 8. *Change in Interest in MOBA Games Following the Study*

Response	n	%
Greatly decreased	2	4
Decreased	11	20
Didn't change	30	56
Increased	9	17
Greatly Increased	2	4

6.5 Sub-Question 3: Explanations vs. Recommendations

What aspects of the AI assistant do players perceive as relevant for improving and reducing their frustration?

A paired t-test indicated that explanations were rated significantly higher than recommendations, $t(53) = 2.72$, $p < .05$, confirmed by Wilcoxon signed-rank test, $Z = -2.56$, $p < .05$. Table 9.

Table 9. *Comparison of Explanation and Recommendation Ratings ($N = 54$)*

Feature	M	SD	Test statistic	p
AI Explanations	3.7	.8	$t(53) = 2.72$	.009
AI Recommendations	3.5	.9		
Wilcoxon confirmation			$Z = -2.56$	.011

Some of the question asked were also open-ended, participants were asked about the best element of the AI Assistant, Table 10 shows that most people preferred the explanations separately.

Table 10. *Frequency of Open-Ended Responses: Most Helpful Aspect and Feature Preference ($N = 54$)*

	Most helpful aspect		Would keep feature	
Response	**n**	**%**	**n**	**%**
Explanations only	29	54	30	56
Item recommendations	13	24	13	24
Both / need both	12	22	11	20

6.6 Sub-Question 4: Itemization Knowledge

Can the AI assistant with explanations influence the players' understanding of the Itemization mechanic?

Table 11 presents knowledge scores before and after the AI condition.

Table 11. *Knowledge Scores Before and After AI Assistance by Experience Level*

Experience	n	Pre-AI M (SD)	Post-AI M (SD)	+/=/−
Never played	28	1.0 (0.8)	1.00 (0.8)	5/19/4
Beginner	15	2.0 (1.0)	3.0 (0.9)	9/6/0
Intermediate	11	3.3 (0.8)	3.8 (0.6)	5/6/0
Total	**54**	**1.7 (1.2)**	**2.1 (1.4)**	**19/31/4**

Note. Scores are out of 4. +/=/− = improved/same/worse.

A Wilcoxon signed-rank test confirmed a significant improvement in overall knowledge accuracy from before to after AI assistance. However, it should be noted that randomness from the participants was involved.

7 Conclusions

This research evaluated an AI itemization assistant for League of Legends that combined a Deep Q-Network for item recommendations with a large language model for real-time explanations. The assistant was assessed by four sub-questions addressing design, retention, feature preferences, and itemization knowledge.

7.1 Main Research Question

Can an AI Itemization Assistant with explanations, for item recommendations in League of Legends, reduce frustration levels and improve player experience in new and novice players.

The AI itemization assistant significantly reduced frustration shown through the negative affect and improved the overall player experience of new and novice League of Legends players.

Negative affect decreased significantly between the unassisted and assisted match ($Z = -2.7$, $p < .05$, $r = .4$), while the rest of the remaining GEQ sub-scales showed even larger improvements: competence ($r = .8$), immersion ($r = .5$), and challenge ($r = .7$), all $p < .05$.

These improvements in competence, immersion, and challenge were consistent regardless of which match was played first, meaning there was no significant order of these effects. Most of the participants had never played the game before, so such an effect can be attributed to general practice and experience gain across the sessions, regardless of the AI assistant.

The Negative Affect results, on the other hand, show a significant order effect. Participants who played without AI first had a decrease of .8 points in the AI match, while the ones that played with the AI first had an increase of .1 in the second match. A general practice effect is also observed in this metric. Participants of the first condition had a great frustration decrease, because it was combining practice with the AI assistant.

In the second condition, the second match had a frustration increase by not having the AI assistant, but it was offset by practice, resulting in a non-significant increase. Taken together, the order effect pattern supports that the AI assistant reduces frustration and is distinct from general practice.

7.2 Sub-Question 1: Design Elements and Frustration Reduction

What design elements and user interaction patterns of the AI assistant are improving the player experience and reduce frustration?

Overall system usability, measured by the SUS, was the strongest single predictor of frustration reduction ($p < .001$), suggesting that ease of use also made sure the user wasn't frustrated.

Interestingly enough the SUS correlation shows that Recommendations were more strongly associated with reducing the negative affect, compared to the Explanations, even though participants reported that the explanations were more useful when asked later. Timing and Disruption were surprisingly non-significant as the system itself required users to look at a separate window.

7.3 Sub-Question 2: AI Assistance, Experience, and Retention

Do players using the AI assistant improve retention levels and does it relate to frustration and experience?

Participants who experienced greater frustration reduction during the AI match were significantly more likely to report an intention to replay the game at a later point.

At the same time, the correlation with recommendation intention did not reach significance ($p = .14$), suggesting that the effect on retention was present but not strong enough to reliably extend beyond the immediate experience.

Overall, 78% of participants reported enjoying the AI match more than the control match (56% slightly better, 22% much better), while 22% rated the two matches as equally enjoyable. Regarding frustration specifically, 54% reported feeling less frustrated during the AI match, while 46% reported the experience as approximately the same.

7.4 Sub-Question 3: Explanations vs. Recommendations

What aspects of the AI assistant do players perceive as relevant for improving and reducing their frustration?

Participants rated the Explanations significantly higher than the Recommendations ($t(53) = 2.72$, $p < .05$, Wilcoxon $Z = -2.56$, $p < .05$).

When asked which element was most helpful, 54% selected explanations, 24% who selected recommendations and 22% found both equally helpful.

Regarding which element participants preferred, 56% chose explanations only, 24% chose recommendations only and the rest wanted both.

The preference indicated confirms the gap identified about the existing available tools, Mobalytics, U.GG, etc, which provide recommendations without context or explanations. Participant clearly prefer the explanations with a goal to understand more, rather than optimize.

7.5 Sub-Question 4: Itemization Knowledge Transfer

Can the AI assistant with explanations influence the players' understanding of the Itemization mechanic?

Overall itemization knowledge improved significantly across the matches ($Z = -2.9$, $p < .05$, $r = .40$), but certain skill groups showed greater improvements than others.

Beginner participants showed the highest improvements, with 9 of 15 improving.

Intermediate participants also improved, from $M = 3.3$ to $M = 3.8$. Most of their questions were correct to begin with, but when observing the experiment, one of the questions appeared obscure enough for intermediate participants to not be sure.

By contrast, participants who had never played League of Legends showed no change. Observing the experiment, it appeared that some of the participants were trying to randomly guess, as some of the questions were outside of their skill level. This was also observed with a minority in the beginner skill bracket.

In the end, AI explanations were most effective for participants who already possessed prior game knowledge and experience to understand and apply the information provided.

8 Discussion

One of the interesting tensions that arose in the results is the contradiction between Sub-questions 1 and 3. In sub-question 1, correlations showed that recommendations were more strongly associated with frustration reduction, but in sub-question 3 participants preferred the explanations.

This can be attributed to the nature of the experiment. During a match, participants needed to know what to buy and the recommendations answer that directly, while explanations can create some decision anxiety. Explanations provided understanding and learning, which requires a slower pace.

An observation during the experiment showed that participants would not read the explanation again in items that they have already seen and read multiple times, thus making them rely on the recommendations more in these cases. As such, the data suggest that recommendations and explanations worked better when paired together.

When it comes to preference, in the study done by Lubos et al.(2024) [19], for a movie recommendation system, reported that 82.5% of the participants preferred the explanations. However, 54% selected explanations as most helpful in this study, which appears low. This figure represents only those who selected explanations exclusively. An additional 22% of participants answered that both features were equally helpful, bringing the overall percentage of those that at least find some use for it at 76%, significantly closer to Lubos's study.

The findings of this study suggest that frustration reflects a broader problem in competitive MOBA games. League of Legends has been live since 2009, and the community has had over 15 years to develop familiarity with the game's mechanics, systems, and meta. Resulting in a community where the average players have a significantly higher level of knowledge compared to a new player attempting to play the game today.

There is too much to learn, which can overwhelm and push the new players out of the flow zone. The results of this study showed that it is possible that an AI assistant can reduce this frustration and nudging the player back to the flow state channel, at least within possibilities with the current features. AI assistants can therefore represent a meaningful step toward making competitive MOBAs more accessible to new players, without removing the complexity of the game's mechanics.

8.1 Future Work

A key priority for future research on expanding this paper, is the separation of the Explanation and Recommendation systems into independent conditions. The current study cannot isolate the two features, since both were delivered together. A design with three conditions: recommendations only, explanations only, and both combined, would allow for more precise measurements of each feature's effect on frustration reduction, learning, and retention.

The experimental setup should also be extended. The current experiment was restricted to a single champion (Sion) in a 1v1 format, which limits generalization. A wider champion pool and in a full 5v5 match environment, alongside a larger training data sample, which could provide a more realistic model and match data, as well as a larger number of participants.

Finally, the assistant itself has great room for development. A future iteration of the LLM could incorporate a dynamic custom explanation system that adapts its language usage and tips to the player's skill level. Alongside this, replacing the separate browser window with an in-game overlay would address most of the usability limitations identified by the SUS score.

8.2 Closing Remarks

This study provides evidence that a real-time AI Itemization assistant can reduce frustration in new and novice League of Legends players, filling a gap that existing tools have left open, which only provide recommendations without explanations. As AI continues to grow, so will the demand of educational tools for players will increase. Professional coaching does exist, but it is unrealistic to be expected that a new player will need one. For general assistants, this research presents a step towards showing the value of AI assistants in high pressure learning environments. And for Games, AI assistants of this kind can develop and adapt alongside the game by scaling to include more game mechanics, which will only improve the player experience further.

References

[1] Csikszentmihalyi, M. *Flow: The Psychology of Optimal Experience* (1990).

[2] Cowley, B., Charles, D., Black, M. & Hickey, R. Toward an understanding of flow in video games. *Comput. Entertain.* **6**, 1–27 (2008). URL https://dl.acm.org/doi/10.1145/1371216.1371223.

[3] Chen, J. Flow in games (and everything else). *Commun. ACM* **50**, 31–34 (2007). URL https://dl.acm.org/doi/10.1145/1232743.1232769.

[4] Sweetser, P. & Wyeth, P. GameFlow: a model for evaluating player enjoyment in games. *Comput. Entertain.* **3**, 3–3 (2005). URL https://dl.acm.org/doi/10.1145/1077246.1077253.

[5] Von Felten, N., Brühlmann, F. & Perrig, S. A. C. Independent Validation of the Video Game Dispositional Flow Scale With League of Legends Players. In *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play*, 44–50 (ACM, Bremen Germany, 2022). URL https://dl.acm.org/doi/10.1145/3505270.3558351.

[6] Jeronimus, B. F. & Laceulle, O. M. Frustration. In Zeigler-Hill, V. & Shackelford, T. K. (eds.) *Encyclopedia of Personality and Individual Differences*, 1–5 (Springer International Publishing, Cham, 2017). URL http://link.springer.com/10.1007/978-3-319-28099-8_815-1.

[7] Kosa, M. & Uysal, A. Need frustration in online video games. *Behaviour & Information Technology* **41**, 2415–2426 (2022). URL https://www.tandfonline.com/doi/full/10.1080/0144929X.2021.1928753.

[8] Khoshnoud, S., Alvarez Igarzábal, F. & Wittmann, M. Peripheral-physiological and neural correlates of the flow experience while playing video games: a comprehensive review. *PeerJ* **8**, e10520 (2020). URL https://peerj.com/articles/10520.

[9] T'ng, S. T., Ho, K. H. & Pau, K. Need Frustration, Gaming Motives, and Internet Gaming Disorder in Mobile Multiplayer Online Battle Arena (MOBA) Games: Through the Lens of Self-Determination Theory. *Int J Ment Health Addiction* **21**, 3821–3841 (2023). URL https://link.springer.com/10.1007/s11469-022-00825-x.

[10] Xia, T., Lin, X., Mo, X., Su, Q. & Ding, S. Players' continuous willingness to play in MOBA game ranking mode: through the lens of self-determination theory and social comparison theory. *Humanit Soc Sci Commun* **11**, 1398 (2024). URL `https://www.nature.com/articles/s41599-024-03934-1`.

[11] Community & Riot-Games. League of legends official wiki. URL `https://wiki.leagueoflegends.com/en-us/`.

[12] Mora-Cantallops, M. & Ángel Sicilia, M. Moba games: A literature review. *Entertainment Computing* **26**, 128–138 (2018). URL `https://www.sciencedirect.com/science/article/pii/S1875952117300149`.

[13] Filipović, A. The role of artificial intelligence in video game development. *Kultura polisa* **20**, 50–67 (2023).

[14] deBeauvais, T., Yee, N. & Ducheneaut, N. Off with their assists: An empirical study of driving skill. In *Proceedings of the 9th International Conference on the Foundations of Digital Games (FDG)* (Society for the Advancement of the Science of Digital Games, 2014). URL `https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/debeauvais-fdg-2014.pdf`.

[15] Warner Bros. Entertainment Inc. Nemesis characters, nemesis forts, social vendettas and followers in computer games. U.S. Patent No. 10,926,179 (2021). United States Patent and Trademark Office.

[16] Tools mentioned: U.gg, op.gg, mobalytics.

[17] do Nascimento Silva, V. & Chaimowicz, L. A Tutor Agent for MOBA Games. *arXiv e-prints* arXiv:1706.02832 (2017). `1706.02832`.

[18] Cunha, R., Machado, M. & Chaimowicz, L. Rtsmate: Towards and advice system for rts games. *Computers in Entertainment* **12** (2014).

[19] Lubos, S., Tran, T. N. T., Felfernig, A., Polat Erdeniz, S. & Le, V.-M. Llm-generated explanations for recommender systems. In *Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization*, UMAP Adjunct '24, 276–285 (Association for Computing Machinery, New York, NY, USA, 2024). URL `https://doi.org/10.1145/3631700.3665185`.

[20] Silva, I., Marinho, L., Said, A. & Willemsen, M. C. Leveraging chatgpt for automated human-centered explanations in recommender systems. In *Proceedings of the 29th International Conference on Intelligent User Interfaces*, IUI '24, 597–608 (Association for Computing Machinery, New York, NY, USA, 2024). URL https://doi.org/10.1145/3640543.3645171.

[21] Kochmar, E. *et al.* Automated personalized feedback improves learning gains in an intelligent tutoring system. In Bittencourt, I. I., Cukurova, M., Muldner, K., Luckin, R. & Millán, E. (eds.) *Artificial Intelligence in Education*, 140–146 (Springer International Publishing, Cham, 2020).

[22] Mnih, V. *et al.* Human-level control through deep reinforcement learning. *Nature* **518**, 529–533 (2015). URL https://doi.org/10.1038/nature14236.

[23] IJsselsteijn, W., {de Kort}, Y. & Poels, K. *The Game Experience Questionnaire* (Technische Universiteit Eindhoven, 2013).

[24] Law, E. L.-C., Brühlmann, F. & Mekler, E. D. Systematic review and validation of the game experience questionnaire (geq) - implications for citation and reporting practice. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '18, 257–270 (Association for Computing Machinery, New York, NY, USA, 2018). URL https://doi.org/10.1145/3242671.3242683.

[25] Brooke, J. Sus: A quick and dirty usability scale. *Usability Eval. Ind.* **189** (1995).

[26] Bangor, A., Kortum, P. T. & Miller, J. T. An empirical evaluation of the system usability scale. *International Journal of Human–Computer Interaction* **24**, 574–594 (2008). URL https://doi.org/10.1080/10447310802205776. https://doi.org/10.1080/10447310802205776.

[27] Nunnally, J. C. *Psychometric Theory* (McGraw-Hill, New York, 1978), 2nd edn.

Appendix

Survey

Block 1: Introduction

Participants were shown the following introductory text:

Welcome to this study on AI assistance in League of Legends Itemization. You will be playing two short matches of the game League of Legends, one without any kind of assistance and one with the AI Assistant for Items. There will be a pre-game survey for you to answer as well as two more after each game. The entire session will take approximately 50–60 minutes. Your responses will help in the understanding of how AI can improve the gaming experience and decrease frustration for new, or novice, MOBA players. Thank you for your participation.

Block 2: Consent Form

Participants selected "Yes, I consent" or "No, I do not consent" in response to the following declaration:

I have read the information letter about the study and understood what it says. I have been able to ask questions, and my questions have been answered to my satisfaction. I have had sufficient time to consider whether I wish to participate. I am participating voluntarily.

Block 3: Demographics

1. **Age** – What is your age? (18–24 / 25–34 / 35–44 / 45–54 / 55+)
2. **Gender** – What is your gender? (Male / Female / Non-binary / Prefer not to say)
3. **Gaming hours** – How many hours per week do you typically play video games? (0–1 / 2–5 / 5–10 / 10–15 / 15+)
4. **Occupation** – What is your current occupation status? (Full-time Student / Part-time Student / Employed / Unemployed / Retired / Unable to work / Prefer not to say)
5. **League of Legends experience** – What would you consider your skill level in League of Legends? (Never played / Beginner / Intermediate / Experienced)

6. **MOBA experience** – Have you played other MOBA games before? (Yes / No)

7. **Platform** – What is your preferred platform to play video games? (PC / Console / Mobile console / Smartphone / Other)

Block 4: Itemization Knowledge (Pre-game)

Participants were instructed: *Please answer the following questions about League of Legends Itemization to the best of your ability. Don't worry if you don't know.*

1. You're playing Malphite (a tank champion) in the top lane against Tryndamere (an attack damage dealer who relies heavily on auto-attacks and critical strikes). Which one of these items would you buy as a second item to counter him? *[4 image options]*

2. You're playing Camille (a fighter with high mobility and true damage) against Aatrox (a sustain-heavy bruiser with strong healing and area-of-effect damage) in top lane. You are one item away from full build and Aatrox is healing back a lot during team fights. What item are you buying? *[4 image options]*

3. You are on low health, after defeating your enemy, and in need of going back to base. You have 1000 gold, but need 100 more to buy a component of an item. What would you do? (Base immediately / Base but not buy the full component / Try to get another wave to buy the item / Wait in base for 100 gold)

4. Are boots better to rush in a losing lane? (Yes / No)

5. What item should you prioritize against Zed (an assassin who deals high burst physical damage) as a mage? *[4 image options]*

6. Which boots are the best option against heavy CC (crowd control)? *[4 image options]*

Participants were then instructed to play Match 1 before continuing.

Block 5: Post-Game Survey — Match 1

All items were rated on a 5-point scale: Not at all (1) – Slightly (2) – Moderately (3) – Fairly (4) – Extremely (5).

GEQ Negative Affect subscale:

1. I felt annoyed
2. I felt bored
3. I felt frustrated
4. I felt irritable

GEQ Competence subscale:

1. I felt competent
2. I felt skillful
3. I felt successful
4. I was good at it
5. I felt capable

GEQ Immersion subscale:

1. I forgot everything around me
2. I was fully occupied with the game
3. I felt completely absorbed
4. I lost track of time

GEQ Challenge subscale:

1. I felt challenged
2. I had to put a lot of effort into it
3. I felt pressured
4. I thought it was hard
5. I felt time pressure

Participants were then instructed to play Match 2 before continuing.

Block 6: Post-Game Survey — Match 2

Identical GEQ items as Block 5, applied to Match 2.

Block 7: Itemization Knowledge (Post-game)

Identical questions to Block 4, with new question IDs, administered after both matches.

Block 8: Retention and Overall Experience

1. Which version of the game did you play first? (With AI / Without AI)
2. How likely are you to play League of Legends again? (Extremely unlikely / Somewhat unlikely / Neither likely nor unlikely / Somewhat likely / Extremely likely)
3. Did this experience increase or decrease your interest in MOBA games? (Greatly decreased / Decreased / Didn't change / Increased / Greatly Increased)
4. Would you recommend League of Legends to a friend new to MOBAs? (Definitely Not / No / Maybe / Yes / Definitely Yes)
5. Which match did you enjoy more? (Match without AI much better / Match without AI slightly better / About the same / Match with AI slightly better / Match with AI much better)
6. Which match did you feel less frustrated during? (Match without AI much less frustrated / Match without AI slightly less / About the same / Match with AI slightly less / Match with AI much less frustrated)

Block 9: System Usability Scale (Treatment Group Only)

All items rated on a 5-point scale: Strongly Disagree (1) – Disagree (2) – Neutral (3) – Agree (4) – Strongly Agree (5).

1. I think that I would like to use this AI assistant frequently
2. I found the AI assistant unnecessarily complex
3. I thought the AI assistant was easy to use
4. I think that I would need the support of a technical person to be able to use this AI assistant
5. I found the various functions in the AI assistant were well integrated
6. I thought there was too much inconsistency in the AI assistant

7. I would imagine that most people would learn to use this AI assistant very quickly
8. I found the AI assistant very cumbersome to use
9. I felt very confident using the AI assistant
10. I needed to learn a lot of things before I could get going with this AI assistant

Block 10: AI Assistant Evaluation (Treatment Group Only)

All items rated on a 5-point scale: Strongly Disagree (1) – Strongly Agree (5).

Explanations:

1. The explanations provided by the AI assistant were easy to understand
2. The explanations helped me learn about itemization in League of Legends
3. The explanations were clear and well-written
4. I could apply what I learned from the explanations to my decision-making

Recommendations:

1. The AI assistant's item recommendations were helpful
2. I trust the AI assistant's item recommendations
3. The recommendations helped me make better itemization decisions

Timing & Disruption:

1. The AI assistant provided recommendations at appropriate times
2. The AI assistant's overlay did not disrupt my gameplay
3. I felt the AI assistant was easy to use during the game

Perceived Impact:

1. The AI assistant improved my overall gameplay performance
2. The AI assistant reduced my frustration about itemization decisions

3. The AI assistant made me feel more confident in my item choices
4. I would use this AI assistant again if it were available

Feature Preference:

1. Which aspect of the AI assistant was most helpful to you? (The item recommendations / The explanations / Both were equally helpful / Neither was particularly helpful)
2. If you could only keep ONE feature, which would it be? (Item recommendations only / Explanations only / I need both features together)