Enhancing Collaboration in Overcooked: Improving AI-Agent Adaptability to Human Player Preferences

A report submitted for the course
COMP3770, Project Work in Computing

12 pt research project, S1/S2 2023

By:
Yuechen Liu

Supervisor:
Dr. Penny Kyburz

Australian National University

School of Computing
College of Engineering, Computing and Cybernetics (CECC)
The Australian National University

October 2023

Declaration:

I declare that this work:

- upholds the principles of academic integrity, as defined in the University Academic Misconduct Rules;
- is original, except where collaboration (for example group work) has been authorised in writing by the course convener in the class summary and/or Wattle site;
- is produced for the purposes of this assessment task and has not been submitted for assessment in any other context, except where authorised in writing by the course convener;
- gives appropriate acknowledgement of the ideas, scholarship and intellectual property of others insofar as these have been used;
- in no part involves copying, cheating, collusion, fabrication, plagiarism or recycling.

October, Yuechen Liu

Acknowledgements

I would like to thank my supervisor, Dr. Penny Kyburz, for her guidance and academic support throughout this research project. My sincere appreciation goes out to all the participants who generously dedicated their time and provided insightful feedback, enriching the quality of my evaluation. I also want to extend my thanks to my friends who stood by me, offering emotional support during challenging times.

Abstract

Collaborative video games, where players work together to achieve common goals, have gained immense popularity in the video game world. However, Training artificial intelligence to act as teammates in such games presents unique challenges. Most AI players are trained to maximise scores, but this doesn't always align with human player preferences. In this research, we delve into the game "Overcooked" to enhance the collaborative capabilities of AI players. Using feedback from human players, we adjusted the AI's decision-making process to better resonate with human preferences. Our evaluation, involving a series of gameplay sessions, revealed that our adjusted AI not only performed slightly better in terms of scores but also aligned more closely with player preferences. This research offers a fresh perspective on designing AI players that harmoniously collaborate with human players.

Table of Contents

Chapter 1

Introduction

The integration of AI agents in collaborative video games has been a popular research topic these years. The primary objective is to ensure that these AI agents not only understand the game mechanics but also align their actions with the preferences of human players. The game "Overcooked" serves as an ideal platform for this exploration, where players collaborate in a confined kitchen space to swiftly prepare and serve dishes to customers.

In recent studies, AI game players are trained using state-of-the-art techniques that involve repetitive engagement with the game simulator Lou et al. (2023). While these techniques optimise gameplay scores, they often overlook the nuances of human player preferences Wang et al. (2021). This is primarily due to the high costs associated with collecting human data during training. As a result, while these AI agents might excel in terms of scores, they frequently misalign with human player strategies, leading to a disjointed gameplay experience Khadpe et al. (2020).

This research is inspired by traditional training methodologies and the lack of lacking human data for the model. Instead of only focusing on score optimisation, the study aims to refine the decision-making phase of the trained AI model to resonate with human player inclinations. The overarching goal is to craft an AI teammate who not only understands the game's mechanics but also tailors its actions to foster a harmonious gameplay experience.

For this project, a simulator of the game "Overcooked" was employed, and the performance of the state-of-the-art AI players Lou et al. (2023) was evaluated based on their collaboration with human players. The primary contribution of this research is the development and evaluation of an AI agent that aligns its gameplay strategies with human player preferences. This alignment is achieved by adjusting the decision-making phase of the trained model, ensuring that the AI agent not only understands the game's mechanics but also resonates with the unique playing preferences of its human teammates.

The structure of this paper is as follows: Chapter 2 provides a brief introduction to the technical contents involved in this research. Hence, Chapter 3 provides a comprehensive review of the literature on AI in collaborative gaming and player experience evaluation. Chapter 4 presents the research methodology, detailing the experiment setup and the adjustments made to the decision-making phase of the trained model. Chapter 5 and Chapter 6 offer a detailed evaluation of the AI agent's performance and relevant discussions, followed by concluding remarks and potential avenues for future research in Chapter 7.

Chapter 2

Background

2.1 Ego Agent

In the context of multi-agent systems and robotics, the term "ego" refers to the primary agent under consideration within a given environment or scenario Hoshino et al. (2004). Specifically, "ego" denotes the agent whose actions, decisions, and perspective are the focal point of analysis or simulation. This usage helps differentiate the main agent from other agents or entities present in the shared space or environment. When multiple agents collaborate or interact, it becomes essential to single out a specific agent's behaviour or viewpoint for detailed study or observation. This agent, whose behaviour or perspective is of primary interest, is termed the "ego" agent. The concept is particularly useful in discussions and analyses related to training or evaluating agents in complex environments, ensuring clarity when referring to the agent of primary interest amidst multiple collaborating or interacting entities.

2.2 Context-Aware Coordination

Context-aware coordination is an approach where the system or agent recognizes and understands the context or environment in which it operates and adjusts its actions or responses accordingly. In the realm of human-AI coordination, this means that the AI system recognizes the behaviour, skill level, or other contextual factors of its human partner and tailors its actions to best complement or coordinate with the human Wang et al. (2022).

2.3 Permutation-Invariant Function

The concept of permutation-invariance is pivotal in the realm of machine learning tasks that deal with set-structured data. A function is deemed permutation-invariant if its output remains consistent regardless of the order of its inputs. In mathematical terms, for a function f, if

$$f(x_1, x_2, \ldots, x_n) = f(x_2, x_1, \ldots, x_n)$$

for all permutations of its inputs, then f is permutation-invariant. Zhou et al. (2022)

2.4 Zero-Shot Evaluation

Zero-shot evaluation is a concept where a model or system is assessed based on its performance in scenarios it hasn't encountered during its training phase. In the realm of human-AI collaboration, this means that the AI is expected to coordinate seamlessly with human partners, even if it hasn't been trained using data from those specific individuals Strouse et al. (2022)

Chapter 3

Related Work

3.1 Human-AI Collaboration in Gaming

In the realm of gaming, the introduction of artificial intelligence (AI) has revolutionised gameplay dynamics, allowing for more complex, challenging, and adaptive environments. While AI's ability to enhance game mechanics and difficulty is evident, its application has evolved to focus on collaboration with human players. AI agents not only challenge players but also understand, adapt to, and cooperate with them, fostering a richer and more engaging gaming environment.

3.1.1 Evolution of AI in Gaming

The landscape of gaming has undergone significant changes with the integration of AI. A primary shift has been the transition of AI from a mere opponent to a collaborative partner. This collaboration requires an intricate understanding not only of game mechanics but also of human intentions, strategies, and preferences. Wang et al. (2020)

The study conducted by Wang et al. mapped the evolving trajectory of AI in gaming Wang et al. (2020). The approach was centred around a comparative examination of AI's roles from past gaming iterations to contemporary collaborative platforms. With a blend of gameplay analysis and player feedback, it was identified that while several AI systems excel in-game mechanics, they often face hurdles when interpreting the subtleties of human intentions and strategies. Such observations highlighted a prevalent gap in AI's ability to anticipate human players, sometimes leading to disjointed collaborations that hamper the immersive experience.

Building on this, another exploration investigated the intricacies of social perceptions within gaming contexts. By combining gameplay analysis, AI behaviour tracking, and player feedback, the study aimed to investigate how AI deciphers nuanced human implications. Lou et al. (2023) The findings were illuminating: AI agents, despite intensive training, often misinterpret or overlook minor indications from human players. This mismatch, due to AI's fixed training methods and lack of flexibility, can lead to unpredictable actions in the game, making it less enjoyable for the player.

3.1.2 Social Perception to Collaborative AIs

The study conducted by Ashktorab et al. highlighted that moments shared between human players, whether they're victories or challenges, are crucial for their enjoyment Ashktorab et al. (2020a). In their experiment, when participants believe they are playing with a human partner instead of an AI, they perceive their partner as more intelligent, creative, and likable, even if the AI behaviour remains consistent. This suggests a potential bias against AI in terms of social perception. However, this bias doesn't impact the actual outcomes of the collaboration. Providing transparency, such as showing confidence scores, can improve social perceptions of AI agents. The findings offer insights for future research on ability improvement and transparency for human-AI collaboration enhancements.

3.2 Overcooked

"Overcooked" is a cooperative multiplayer cooking simulation game developed by Ghost Town Games. In the game, players collaborate to prepare and serve a variety of dishes to customers within a specified time frame. The game is set in dynamically changing kitchens that present various challenges, from moving platforms to obstacles like fire and ice, making the gameplay both engaging and challenging Games (2016)

3.2.1 Collaboration in Overcooked

"Overcooked" presents a dynamic environment that requires players to continuously adapt their strategies in response to evolving challenges. The game's design, characterised by varying kitchen layouts, unpredictable obstacles, and diverse dish demands, ensures that players cannot remain static in their approach. In this case, the game promotes collaboration, as players are compelled to coordinate their actions and communicate consistently. Ltd. (2016)

The study conducted by Rosero et al. provides insights into the intricacies of team communication within the game. One of the key findings was the importance of effective communication for successful gameplay. Players who communicated their intentions and actions clearly and timely were more likely to succeed in the game's challenges. This is especially true in levels where the kitchen layout is complex, and players need to coordinate their movements to avoid collisions and ensure smooth gameplay.Rosero et al. (2021)

Furthermore, the research highlighted how players' perceptions of their teammates influenced their collaborative strategies. When players perceived their teammates, not matter human or AI, as competent and reliable, they were more likely to delegate tasks and trust them with crucial responsibilities. On the other hand, if a player perceived their teammate as unreliable, they might take on more tasks themselves, leading to potential inefficiencies and bottlenecks.

3.2.2 Overcooked Simulator

The Overcooked Simulator is a virtual representation of the cooperative cooking game “Overcooked.” The simulator offers a simplified version of the Overcooked game environment. Carroll et al. (2020) The primary objects in this version are onions, dishes, and soups. The gameplay involves a series of high-level actions: players must place three onions in a pot, allow them to cook for 20 timesteps, transfer the cooked soup into a dish, and then serve it. Successfully serving a dish rewards all players with a score of 20. Players can choose from six possible actions: moving up, down, left, or right, remaining stationary, and “interact”. The “interact” action is context-sensitive, meaning its effect depends on the tile the player is facing. For instance, it could involve placing an onion on a counter. The game layouts feature one or more onion dispensers and dish dispensers, ensuring an endless supply of essential ingredients and tools.

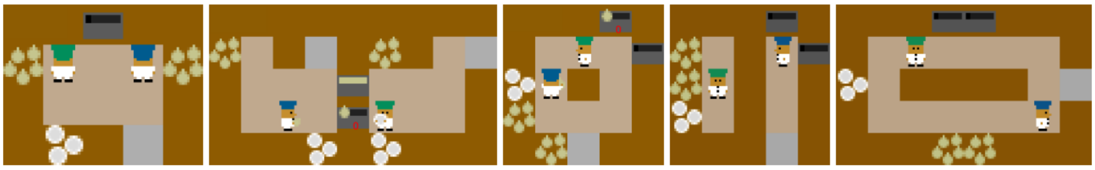

Figure 3.1: Experiment layout Rosero et al. (2021)

The layouts of the simulator are designed to test different aspects of collaboration, coordination, and strategy. From left to right in Figure 3.1 are the illustrations of the five different layouts. The Cramped Room layout emphasises basic coordination challenges due to its limited space, making player collisions highly probable. Asymmetric Advantages evaluate players’ ability to adopt advanced strategies that capitalise on individual strengths, which indicates that little collaboration is involved in this scenario. In the Coordination Ring layout, players are tasked with synchronising their movements to navigate between two pivotal points on the map. Conversely, Forced Coordination eliminates the issues of movement collisions, compelling players to formulate a collaborative strategy, given that a single player cannot complete a dish independently. Lastly, the Counter Circuit layout introduces an unconventional coordination approach where players transfer onions across the counter to the cooking pot, rather than the typical method of walking them around.Rosero et al. (2021)

3.3 Predictive Entropy-Conditioned Agent

The Predictive Entropy-Conditioned Agent (PECAN) is a state-of-the-art AI framework for gaming scenarios, particularly those requiring intricate human-AI collaborations. Lou et al. (2023) Using advanced predictive algorithms, PECAN is structured to interpret and anticipate in-game dynamics, ensuring its actions and decisions are consistent with human player strategies. This robust alignment is central to PECAN's design philosophy, aiming to bridge the prevalent gaps in conventional AI gaming agents and to provide a more cohesive, efficient, and rewarding collaborative gameplay experience.

3.3.1 Permutation-Invariant Function in PECAN

For PECAN, the permutation-invariant property is indispensable for the context encoder module. This module is tasked with predicting a partner's level-based context derived from past trajectories. By ensuring that the function's output remains consistent irrespective of the order of past trajectories, PECAN can make more robust and reliable predictions. This design choice is crucial as the sequence of transitions within a trajectory and the trajectories themselves are irrelevant to the partner's level. Zhou et al. (2022) Therefore, ensuring permutation invariance allows PECAN to handle diverse and unpredictable human behaviours effectively, making it adaptable and robust in real-world scenarios where data order can be arbitrary and unpredictable. Lou et al. (2023)

3.3.2 Context-Aware Coordination in PECAN

PECAN introduces a unique approach to context-aware coordination by modelling context identification as a trajectory-to-class mapping. In this setup, contexts are visualized as class labels, which indicate the corresponding group—be it low, medium, or high-level partners—based on their training duration. The context encoder in PECAN serves as a classifier that predicts the context from past trajectories. This design choice ensures that the ego agent's responses are tailored to the inferred skill level of the partner Lou et al. (2023)

By conditioning on the predicted context $\hat{c}$ indicating the partner's level, the ego agent's policy $\pi_{\text{ego}}(\cdot|s, \hat{c})$ learns level-specific coordination skills. This is more universal than previous partner-specific coordination skills. As a result, PECAN achieves a level-based common best response (BR) to partners of different levels. This approach distinguishes PECAN from other methods, such as posterior sampling, which view context identification as a distribution. In PECAN, it's modelled as a mapping from trajectories to classes Lou et al. (2023).

The significance of context-aware coordination in PECAN is underscored by its ability to conduct zero-shot evaluations with unfamiliar partners. As the ego agent accumulates more trajectories, its belief system refines, leading to increasingly accurate predictions. This ensures that even when faced with novel partners, PECAN can adapt and coordinate effectively without the need for additional training. Empirical validation in the PECAN paper further highlights the adaptiveness of the system, showcasing the policy ensemble's role in enhancing partner diversity and the context-awareness of the ego agent's policy Lou et al. (2023).

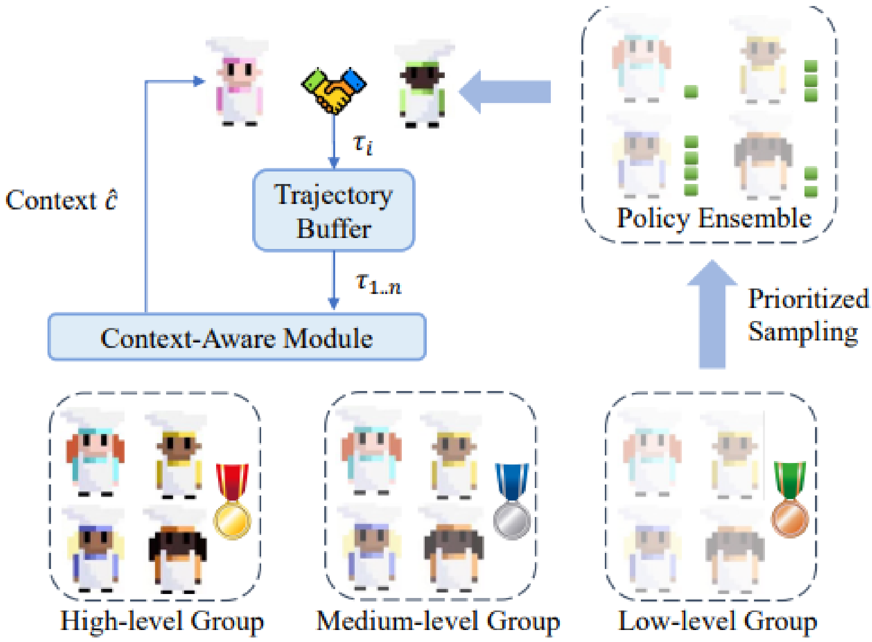

Figure 3.2: The ego agent collects trajectories during collaborating with the partner and recognise the partner's level-based context at the beginning of each episode with a pretrained context-aware module. Adapted from PECAN: Leveraging Policy Ensemble for Context-Aware Zero-Shot Human-AI Coordination. Lou et al. (2023)

3.3.3 Zero-Shot Evaluation

When introduced to unfamiliar partners, PECAN relies on analyzing collected trajectories to deduce the context of the partner. Mathematically, the context encoder can be represented as $\hat{c} = f(T)$, where $\hat{c}$ is the predicted context, T represents the past trajectories, and f is the function (context encoder) that maps trajectories to contexts. As it gathers more trajectories, PECAN refines its belief system, leading to enhanced prediction accuracy. Strouse et al. (2022) This methodology sets PECAN apart from other systems that might necessitate explicit training with each new partner's data. Instead, PECAN's design allows it to adapt and coordinate based on the deduced context without undergoing additional training sessions.

```
function ContextIdentification(T):
    Initialize context encoder f
    for each trajectory t in T:
        context_hat = f(t)
    end for
    return context_hat
end function
```

During the zero-shot evaluation with a novel partner, PECAN uses the collected trajectories to infer the context of the partner. As more trajectories are collected, the ego agent's belief narrows, and the prediction becomes more accurate. Importantly, even though the context inference uses past trajectories, PECAN does not fine-tune or update any parameters during evaluation, ensuring that the evaluation with a novel partner remains zero-shot. Lou et al. (2023)

This approach allows PECAN to adapt to new partners without the need for additional training, making it particularly suited for real-world scenarios where the AI system might encounter a wide range of human partners with varying behaviours and skill levels.

3.3.4 Limitations

Although PECAN is demonstrating optimising results in getting AI to work collaboratively with humans, it still has some limitations. One of the main hurdles is how PECAN tries to decipher human behaviours. It primarily observes patterns in gameplay actions, but this might not fully capture the subtle strategies or unique approaches each player adopts. Moreover, PECAN's training doesn't incorporate real-time feedback from human players, mainly because gathering such data is both time-consuming and challenging. Lou et al. (2023)

To improve AI-human collaboration in gameplay, understanding human feedback after initial training is essential. A post-training preference learning module could be a possible improvement. Using feedback from human players, the AI's decision-making can be adjusted to better match player preferences. While human data isn't used in the initial training, this post-training adjustment ensures the AI aligns with human tendencies. Adachi and Hodson (2018) This approach maintains the benefits of zero-shot coordination, as in PECAN, but also makes the AI more adaptable to specific human feedback, leading to better gameplay coordination.

3.4 Player Gaming Experience Evaluation

It is important to have an accurate measurement of the player experience in video games, and having a valid and credible measuring metrics is fundamental to the understanding of the impact of video games, designing and developing video games, and also, effectively integrating AIs into games to enhance the player experience while not breaking the balance of the game. In this section, some commonly used evaluation metrics are introduced and evaluated.

3.4.1 Player Experience of Need Satisfaction

The Player Experience of Need Satisfaction (PENS) is a specialized tool designed to assess the psychological needs of players during gameplay. Rooted in the Self-Determination Theory (SDT), PENS focuses on three core intrinsic needs: autonomy, competence, and relatedness Johnson et al. (2018).

- **Autonomy** The autonomy metric refers to the evaluation of the sense of agency and volition a player feels while interacting with a game. In video games, this sense of autonomy is often manifested when players are presented with meaningful choices that have tangible impacts on the game's outcome or narrative Hopson (2001). When evaluating autonomy, PENS focuses on players' feelings of freedom within the game. For example in strategic games, the sense of freedom should come from allowing players to decide on a particular approach or tactic. It gauges whether players feel they have meaningful choices and whether they feel forced into particular decisions Emmerich and Masuch (2016). A game that successfully fosters autonomy ensures that players feel in control and that their choices carry weight.

- **Competence** Competence is the evaluation of whether the players feel like they are effective and capable in their interactions with the game Hopson (2001). This feeling of competence is often derived from moments where players overcome challenges, achieve set goals, or progress through the game's levels or narratives. The crucial role for boosting the competence of players is the feedback and rewards system. When players receive clear and constructive feedback on their actions, it can significantly enhance their sense of competence. PENS evaluates this by evaluating players' feelings of mastery and achievement, and by assessing whether the game's challenge level aligns with the player's skill, ensuring neither excessive ease nor overwhelming difficulty.

- **Relatedness** Relatedness addresses the player's need to connect with others and to feel a sense of belonging. In multiplayer or team-based games, this is often about the quality and depth of interactions with other players.Johnson et al. (2018) In single-player games or games that involve human-AI collaborations, a sense of relatedness can emerge through meaningful interactions with the non-player characters. When assessing relatedness, PENS seeks to understand players' feelings of connection to other characters or players and their sense of belonging within the game's universe. A game that demonstrates strong relatedness ensures that players feel they are a meaningful part of the game's community or story.

3.4.2 Game Experience Questionnaire

The Game Experience Questionnaire (GEQ) is crafted to quantitatively analyse the multifaceted experiences of players as they interact with video games. Johnson et al. (2018) In this questionnaire, the game experience is gathered as scores on seven components: Immersion, Flow, Competence, Positive and Negative Affect, Tension, and Challenge IJsselsteijn et al. (2013).

- **Immersion** Immersion is the matrix that delves deep into the player's sense of being engrossed throughout the game experience. An immersive game not only contains good visual and auditory aspects but also the sensation of being transported to another realm. Immersion captures the essence of players feeling disconnected from their immediate surroundings, becoming a part of the game's narrative and environment. In a collaborative game, immersion is often related to the compelling storyline and character settings, and appropriate challenges and tasks settings. These factors collectively draw the player into the game's world Fox and Brockmyer (2013).

- **Flow** Flow is a nuanced concept that originates from positive psychology, representing a state where individuals feel completely engaged in an activity, experiencing a balance between their skills and the challenge at hand Csikszentmihalyi (1990). In the context of multiplayer games, flow is closely tied to the synergy between players. Achieving flow becomes a collective endeavour, where the challenge presented by the computer-controlled elements must be matched by the combined skills of the team. Moments of flow are most profound when the game's AI challenges are neither too easy nor too hard, requiring players to collaborate effectively and utilize each member's strengths Voiskounsky et al. (2004).
- **Tension** Tension in the GEQ is multifaceted. While it can arise from challenging gameplay elements, it also encompasses the emotional highs and lows players experience due to narrative elements, unexpected in-game events, or competitive scenarios. Tension can be both positive, adding to the excitement and engagement, or negative, leading to feelings of frustration or anxiety Ravaja et al. (2006). In cooperative multiplayer games, tension often arises from the pressure to perform well for the team and the unpredictability of challenges. Since players rely on each other to navigate challenges, the tension escalates when the team faces high-risk scenarios or near defeats. However, the shared experience of overcoming these tense moments can lead to heightened satisfaction IJsselsteijn et al. (2013).
- **Competence** Competence is the player's perception of their abilities in the game, which is also a metric evaluated for PENS. In multiplayer games, players derive a sense of competence from contributing to team successes. Recognizing when to lead, when to support, and when to adapt based on the team's needs and the challenges is crucial Ryan et al. (2006).
- **Positive and Negative Affect** Positive Affect and Negative Affect in cooperative multiplayer games are influenced by team dynamics. hared victories, effective collaboration, and moments of collective brilliance can amplify positive emotions. Conversely, missteps, miscommunications, or perceived shortcomings in team strategy can lead to negative emotions. While Positive Affect gauges the elation, thrill, and pleasure derived from gameplay, Negative Affect captures feelings of irritation, disappointment, or even anger. These metrics directly influence a player's overall perception of the game and their desire to continue playing Sherry et al. (2006).
- **Game Value** Game Value evaluates the overall worth of the game, and it is a reflection of whether players perceive the game as a worthy investment of their time and effort IJsselsteijn et al. (2013). In cooperative multiplayer games, game values are often evaluated based on the quality of team interactions, the challenge

presented in the game, and the opportunities for meaningful collaboration. Players value games that foster teamwork, present engaging AI challenges, and offer opportunities for shared achievements and memories. The sense of accomplishment derived from collaborating to overcome challenges can significantly enhance the perceived value of the game.

3.4.3 Human-AI Collaboration Games Evaluation

In recent years, the evaluation of player experience in human-AI collaborative games has gained significant attention. the evaluation of player experience in human-AI collaborative games is multifaceted, encompassing gameplay outcomes, social dynamics, and player perceptions.

One potential measurement metrics is to include factors such as trust, mutual understanding, and shared decision-making, which are important in assessing the quality of collaboration. Furthermore, the game outcomes, efficiency of task completion, the number of objectives achieved, and the overall game score are also measured. Wang et al. (2020) This could be done by encompassing post-game surveys to gauge player satisfaction, and then analyse the player's overall experience and feelings towards the AI agent. Additionally, the AI agent's behaviour, focusing on its responsiveness, adaptability, and effectiveness in the game can also be surveryed and analysed. The communication patterns, which entail the frequency and type of communication between the human player and the AI agent, is also a vital component to understand the depth of the collaboration.

Another metric is designed to measure how effectively human players and AI agents could work together to achieve in-game objectives. Their study also delved into understanding the AI agent's contribution, assessing its role and impact on the game's success. Through surveys, they captured players' perceptions, emphasizing how players viewed the AI's collaborative abilities and its role in enhancing or hindering the gameplay experience Ashktorab et al. (2020b).

Chapter 4

Experiment

This chapter delves into the user study carried out to assess the preferences of human players when interacting with two distinct AI models: the original and the updated versions, within the context of the "Overcooked" game simulator. All pertinent documents and materials related to this experiment can be found in Appendix B.

4.1 Target Participant

The target participant group for the research project included individuals who had experience playing the video game Overcooked. These players were expected to be familiar with the game mechanics and have a reasonable understanding of the cooperative nature of gameplay. Adults from any professional background were invited to participate. Posters were displayed around ANU detailing the information about the target participant group. These posters also provided the researcher's phone number and email address, allowing potential participants to reach out and inquire about specific details regarding the research.

The proposed number of participants was 20. The participants for the research project were carefully selected based on their relevant experience and familiarity with the video game Overcooked. By targeting individuals who had played the game, the research aimed to benefit from their insights and opinions related to the game mechanics and the cooperative nature of gameplay. Their firsthand experiences in Overcooked were expected to allow them to provide informed feedback on AI teammates and their preferences for collaborative gameplay.

4.2 Experiment Setup

The experiment was set up using the previously introduced simulator, which was integrated with the trained PECAN model. In the original agent configuration, the policy, denoted as $\pi_{\text{ego}}(\cdot|s, \hat{c})$, was directly responsible for producing an action based on its output. An action was sampled directly from this policy. However, in the updated model, an additional layer of refinement was introduced. The heuristic of the game, which computes the cost for the agent to conduct each action, is slightly adjusted based on the participant's preference for each potential action provided in the feedback. This adjustment to the heuristic effectively adds an adjustment layer before the action sampling process.

It's crucial to note that this adjustment was not designed to drastically modify the agent's decisions. Instead, its primary function was to subtly influence the agent's choices, especially in situations where multiple actions had comparable probabilities in the policy. This nuanced adjustment ensured that the AI's decisions were more in line with the observed preferences of the participant, thereby aiming to enhance the collaborative gameplay experience.

4.3 Participants Interaction

Participants interacted within this simulated environment using a straightforward control scheme: the 'ASDW' keys were designated for directional movements, while the space bar was employed for special interactions. The nature of these special interactions was context-sensitive, determined by the avatar's current position in the game and the proximate objects available for interaction. This setup was designed to emulate the gameplay mechanics of "Overcooked", ensuring that participants' actions and decisions within the simulator mirrored the challenges and dynamics they would encounter in the actual game.

4.4 Procedure

The experiment took 40 to 60 minutes, and invited participants to experience the "Overcooked" game across five distinct scenarios. It began with a demonstration of the game's dynamics, after which participants shared their gameplay preferences. They then played ten trials, interacting with two versions of in-game characters, with the order varied for unpredictability. After each scenario, feedback was gathered on the gameplay and character interactions. The session concluded with participants offering suggestions to improve the overall game experience.

The following sequence describes the list of steps taken in conducting the experiment:

1. Upon arrival, participants were given a brief introduction about the study's focus on enhancing collaboration between AI and human players in "Overcooked".

They were then presented with an electronic consent form to sign, indicating their voluntary participation and understanding of the data collection process.

2. Before diving into the main experiment, participants were shown a demonstration of how the AI behaves when interacting with another AI in "Overcooked". They were encouraged to observe the AI's actions closely, especially noting any behaviors they found undesirable.

3. Based on the feedback from the demonstration, an updated model was set up by adding an adjustment layer to modify the AI's actions before they were executed in the game.

4. Participants were then seated in front of a computer setup, equipped with the "Overcooked" game simulator. Here, they would play alongside two different AI models across five distinct scenarios. Each scenario was timed, and participants were encouraged to collaborate with the AI to achieve the best possible outcome.

5. After each interaction with an AI model, participants were prompted to share their immediate feedback and preferences. They were asked to comment on their collaboration experience with the AI and to specify which of the two AI models they preferred for that particular scenario.

6. This process was repeated for all five scenarios, ensuring that participants had ample opportunity to interact with both the original and updated AI models.

7. Once all scenarios were completed, participants were invited for a semi-structured interview. This session aimed to gather deeper insights into their gameplay experience, their preferences between the two AI models, and any suggestions they might have for further improving the AI's behavior in "Overcooked".

8. After the interview, participants were thanked for their invaluable contribution to the research. They were reminded that their feedback would play a crucial role in enhancing the AI's collaboration capabilities in "Overcooked".

9. The session concluded with a debriefing, where participants were once again thanked for their time and participation. Those who had registered through the university's research participation system were credited for their involvement in the study.

4.5 Ethical Clearance

Prior to initiating the experiments, ethical approval was secured. The *National Statement on Ethical Conduct in Human Research* was consulted to ensure adherence to the mandated responsibilities and guidelines for this research endeavour. In alignment with Section 2.2.6 of the *National Statement*, a *Participant Information Sheet* was crafted, detailing the study's objectives, potential risks, benefits, and methodologies. The ethics approval of this project is under Protocol H/2023/1195 - *Enhancing Collaboration in Overcooked: Improving AI-Agent Adaptability to Human Player Preferences.*

The documentation provided with this variation submission can be found in Appendix B.

4.6 Data Collection

In this experiment, both qualitative data and quantitative data were collected for analysis purposes.

The primary objective of the quantitative data collection was to evaluate the performance scores attained by players when interacting with the two AI models in the game. The scores from each game session served as a metric to gauge the competency of the players in different scenarios. This is used to see whether player scored higher in general with the updated model. Additionally, a correlation analysis was employed to discern if there was a statistically significant relationship between the scores and the players' preferences for the AI models. The hypotheses for the correlation analysis were:

- **Null Hypothesis (H0):** There is a significant correlation between player preference and the scores they achieved with each AI model, suggesting that players may prefer the AI model with which they score higher.
- **Research Hypothesis (H1):** There is no significant correlation between player preference and the scores they achieved with each AI model, indicating that player preference is independent of the scores they achieve.

In addition to the quantitative metrics, qualitative data was also collected to delve deeper into the subjective preferences and feedback of the players. After each game session, participants were asked to indicate their preference between the two AI models for that specific scenario. The count of these preferences was then used to test the correlation between player preferences and the scores they achieved in the game. Furthermore, participants were encouraged to provide feedback on their gameplay experience, pointing out any specific AI behaviours that influenced their decisions. This feedback was analysed to understand the relationship between player preference, scores, and their comments both before and after the gameplay sessions.

The interview guide and the detailed questions posed to the participants can be found in Appendix B.

Chapter 5

Evaluation

This chapter presents and evaluates the data collected from the experiment. These data are sampled from 20 participants with past Overcooked game experience. The participants were screened before their participation to ensure that they are eligible to participate. Full result can be reviewed in Appendix A.

5.1 Player Scores Overview

Overall, the players have played 100 rounds with each agent, resulting in a total of 200 rounds played. The scores achieved by the players when paired with each agent are summarized in the table below.

Table 5.1: Summary of Player Scores with Original and Updated Models and T-test Results

	Original Model	**Updated Model**
Mean	201.20	201.80
Standard Deviation	67.16	63.69
Standard Error of Mean	15.02	14.24
T-value	-0.0645	
P-value	0.4743	
Degree of Freedom	198	

From Table 5.1, it can be observed that the mean scores achieved with both the original and updated models are quite similar, with only a slight increase in the mean score for the updated model. The standard deviation for the original model is slightly higher, indicating a wider spread of scores around the mean compared to the updated model. The standard error of mean, which provides an estimate of how much the sample mean

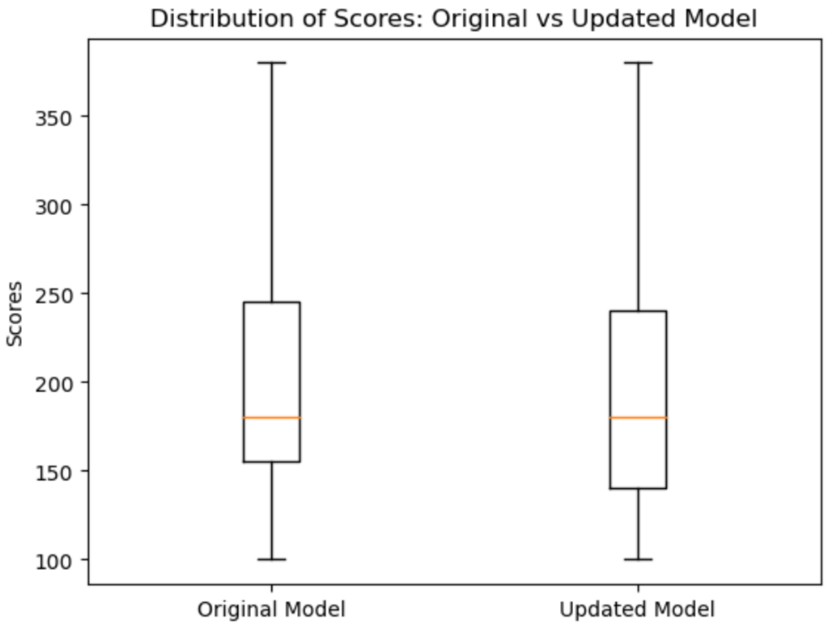

Figure 5.1: Box plot for the distribution of the scores obtained during collaborations with each model

is expected to vary, is also slightly higher for the updated model. The similarity between the means can also be observed from the box plot, as the position of the orange line are approximately the same.

The t-test was conducted to compare the scores of players when playing with the original model versus the updated model. The mean score with the original model was 201.20 with a standard deviation of 67.16, while the mean score with the updated model was 201.80 with a standard deviation of 63.69.

The t-value of -0.0645 indicates a very slight difference in the means of the two groups, with the scores from the original model being slightly lower than those from the updated model. However, this difference is negligible.

The one-tailed p-value of 0.4743 is greater than the common alpha level of 0.05. This means that the observed difference in scores between the two models is not statistically significant. In other words, there is no evidence to suggest that the scores achieved with the updated model are significantly different from those achieved with the original model.

Furthermore, the lack of a significant difference between the scores achieved with each model suggests that the addition of the adjustment layer, which was designed to capture the subjective preferences of the players, did not adversely affect the performance of the model. This is a positive outcome, as it indicates that the model can be tailored to individual player preferences without compromising its overall effectiveness in the game.

In conclusion, based on the t-test results and the observed data, there is no significant difference in player scores when playing with the original model compared to the updated model. This reinforces the notion that integrating subjective player preferences into the AI model, through the adjustment layer, is a viable approach that does not hinder the model's performance.

5.2 Player Preferences Overview

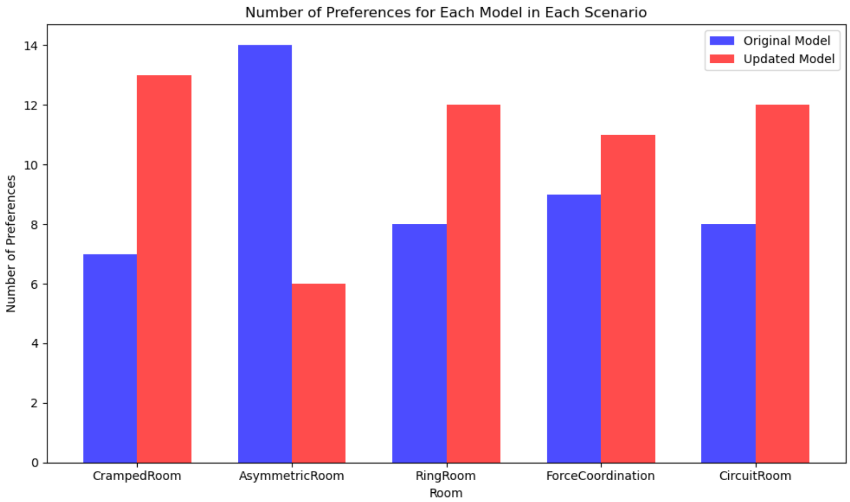

Figure 5.2: Bar chart of Participants Preferences for Original and Updated model in each scenario

Table 5.2: Percentage of Player Preferences for Original and Updated Models

Room	**Original Model (%)**	**Updated Model (%)**
Cramped Room	35.0	65.0
Asymmetric Room	70.0	30.0
Ring Room	40.0	60.0
Force Coordination	45.0	55.0
Circuit Room	40.0	60.0

From the data, it is evident that most rooms have a higher preference for the updated model. This suggests that the adjustments made to the updated model are generally favored by the players, indicating its effectiveness in enhancing the gameplay experience.

The Cramped Room stands out with the highest preference for the updated model, at 65%. A plausible explanation for this is that participants observed the original agent collaborating with the baseline agent in the Cramped Room during the demonstration. As such, the adjustments made for the updated model could be more impactful and noticeable in the same room, leading to a higher preference.

Conversely, the Asymmetric Room shows the lowest preference for the updated model, at only 30%. Given that this room doesn't necessarily emphasize collaboration, one possible reason for this lower preference could be that the updated model is achieving lower scores compared to the original model in this particular room. This hypothesis will be further explored and verified in subsequent analyses.

5.3 Score-Preference Correlation

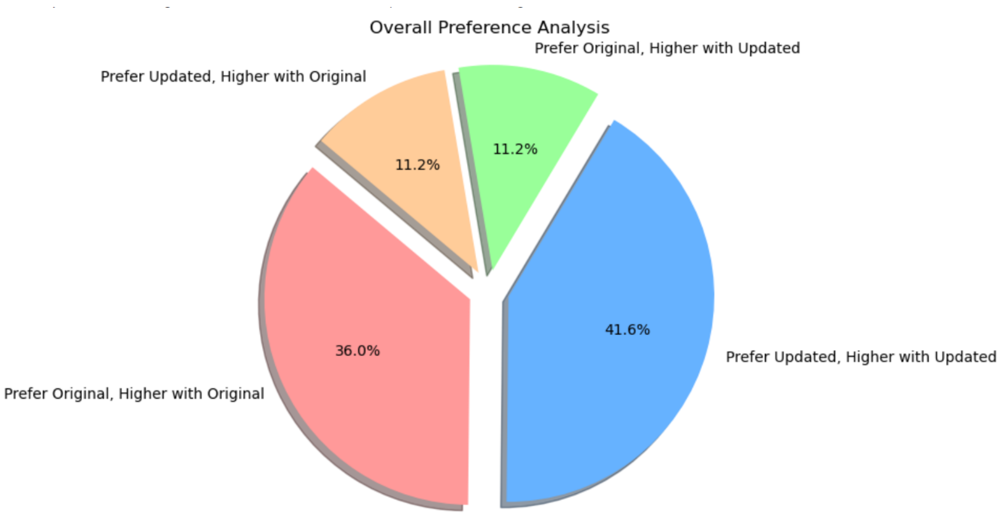

Figure 5.3: Pie Chart for Player Preferences and Score Distribution

Table 5.3: Overall Correlation between Preference and Scores

Metric	**Value**
Overall Chi-square value	24.6800
Overall P-value	0.0000

In the study, player preferences were collected post-gameplay in each layout. For the purpose of analysis, these preferences were numerically encoded: a preference for the original model was denoted as 0, while a preference for the updated model was denoted as 1. Player scores with each model were compared, and the model yielding the higher score was identified. This information was then converted into a numerical label to facilitate statistical analysis.

The chi-square test was employed to determine if there is a significant association between two categorical variables: player preference and the model that yielded a higher score. The null hypothesis (H_0) posits that there is no association between player preference and the model that produced a higher score. Conversely, the research hypothesis (H_1) asserts that there is a significant association between the two.

The results, as presented in Table 5.3, show a significant correlation with a p-value of 0.0000, which is below the common alpha level of 0.05. This indicates that the observed correlation is statistically significant, leading to the rejection of the null hypothesis in favor of the research hypothesis.

From the results, it is evident that there is a strong correlation between player preference and the scores they achieved in each game. This suggests that participants tend to prefer the model that aids them in obtaining a higher score. In other words, the performance of the model, as reflected by the scores, plays a pivotal role in shaping player preferences.

It is worth noting that while this section provides an overall analysis, data specific to each layout were also dissected and analyzed separately in subsequent subsections. For a visual representation of the preference distribution, refer to the pie chart in Figure 5.3.

5.3.1 Layout1: Cramped Room

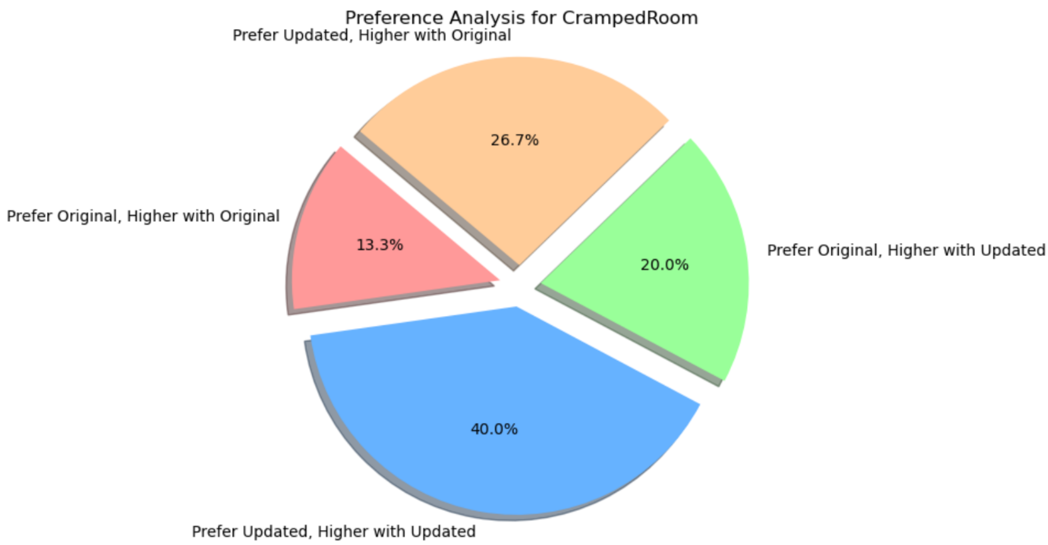

Figure 5.4: Pie Chart for Player Preferences and Score Distribution in Cramped Room Layout

In the Cramped Room layout, the relationship between player preference and the model that yielded a higher score was examined. The null hypothesis (H_0) for this analysis posits that there is no significant association between player preference and the model that produced a higher score. Conversely, the research hypothesis (H_1) suggests that there is a significant association between the two variables.

Table 5.4: Correlation between Preference and Higher Score in Cramped Room

Metric/Combination	**Value**
Prefer original and scored higher with original	2
Prefer updated and scored higher with updated	6
Prefer original and scored higher with updated	3
Prefer updated and scored higher with original	4
Chi-square value	0.0000
P-value	1.0000

The chi-square test, a statistical measure used to determine the independence of two categorical variables, yielded a value of 0.0000. This value, in conjunction with the p-value of 1.0000, indicates that we fail to reject the null hypothesis. In other words, there is no significant correlation between player preference and the model that resulted in a higher score in the Cramped Room layout.

It's worth noting that participants had previously observed the model collaborating with the baseline model in this layout and provided feedback on the actions taken by the original model. This feedback was then used to adjust the model. The lack of correlation between score and preference in this layout could be attributed to the fact that participants' feedback was specifically targeted at this layout, leading the model to opt more for their preferred actions. This could mean that even if the score wasn't necessarily higher, the model's actions were more in line with the participants' expectations or preferences based on their feedback. Additionally, the adjustments made to the model based on feedback might have led to more diverse strategies being employed, which, while not necessarily yielding higher scores, might have been perceived as more engaging or interesting by the participants.

5.3.2 Layout2: Asymmetric Room

Table 5.5: Correlation between Preference and Higher Score in Asymmetric Room

Metric/Combination	**Value**
Prefer original and scored higher with original	14
Prefer updated and scored higher with updated	6
Prefer original and scored higher with updated	0
Prefer updated and scored higher with original	0
Chi-square value	15.5215
P-value	0.0001

In the Asymmetric Room layout, an analysis was conducted to explore the relationship between player preference and the model that yielded a higher score. The null hypothesis for this investigation asserts that there is no significant association between player preference and the model that produced a higher score. On the other hand, the research

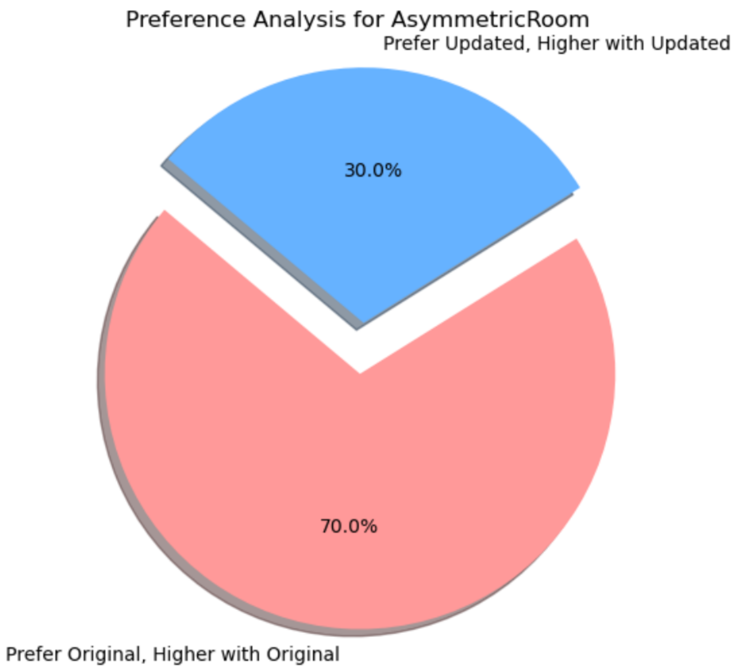

Figure 5.5: Pie Chart for Player Preferences and Score Distribution in Asymmetric Room Layout

hypothesis proposes a significant association between these two variables.

Utilizing the chi-square test, a statistical value of 15.5215 was obtained. Coupled with a p-value of 0.0001, which is well below the common alpha level of 0.05, this indicates a rejection of the null hypothesis in favor of the research hypothesis. Thus, there is a significant correlation between player preference and the model that resulted in a higher score in the Asymmetric Room layout.

The Asymmetric Room is characterized by its lack of necessity for collaboration. In such a scenario, both the model and the participant can engage in the game independently. Given this independence, participants might not have had ample opportunities to provide feedback regarding collaborative actions. Consequently, their preference could be predominantly influenced by the scores they achieved. This suggests that in scenarios where collaboration isn't integral, players might prioritize performance, as reflected by their scores, over other aspects of gameplay. Additionally, the absence of collaborative dynamics might lead to a more straightforward evaluation by the participants, where they lean towards the model that consistently offers them a higher score.

For a comprehensive visual representation of the preference distribution in the Asymmetric Room, readers are directed to the accompanying pie chart. This visualization further elucidates the alignment of preferences and scores, offering insights into participant reactions to the model's performance in this distinct layout.

5.3.3 Layout3: Ring Room

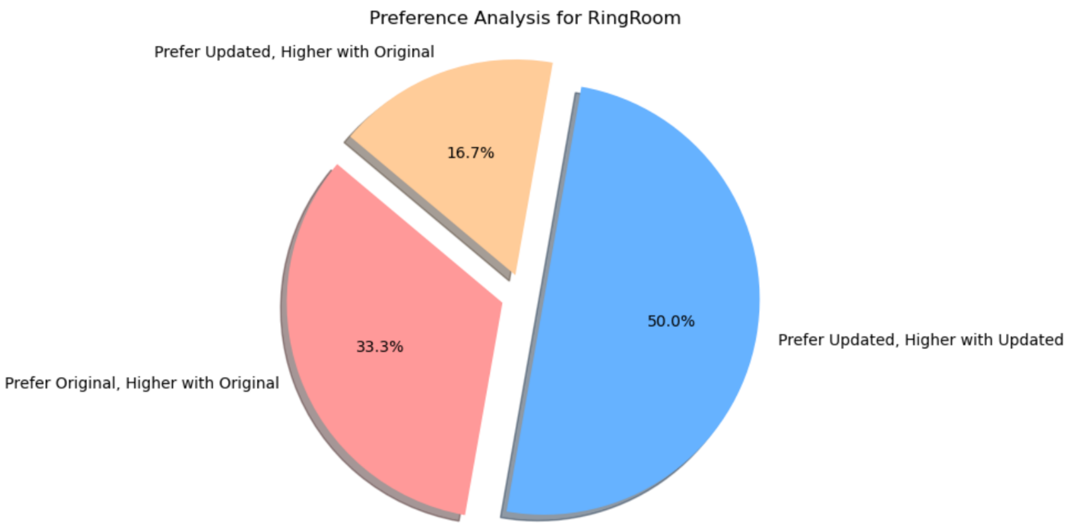

Figure 5.6: Pie Chart for Player Preferences and Score Distribution in Ring Room Layout

Table 5.6: Correlation between Preference and Higher Score in Ring Room

Metric/Combination	**Value**
Prefer original and scored higher with original	6
Prefer updated and scored higher with updated	9
Prefer original and scored higher with updated	0
Prefer updated and scored higher with original	3
Chi-square value	6.2500
P-value	0.0124

In the context of the Ring Room layout, an investigation was undertaken to discern the relationship between player preference and the model that yield a higher score. The null hypothesis for this study posits that there is no significant association between player preference and the model that achieved a higher score. Conversely, the research hypothesis suggests a significant association between these two variables.

Employing the chi-square test, a statistical value of 6.2500 was derived. With a p-value of 0.0124, which falls below the conventional alpha level of 0.05, this leads to the rejection of the null hypothesis in favor of the research hypothesis. As a result, a significant correlation between player preference and the model that yielded a higher score in the Ring Room layout is established.

The unique design of the Ring Room, where participants and the model cannot traverse the space simultaneously, places a heightened emphasis on the model's actions. Specifically, if the model obstructs the path, it can detrimentally impact the score and significantly mar the player experience. This direct influence of the model's actions on both the score and player satisfaction provides a plausible explanation for the observed correlation between preference and score.

Moreover, the Ring Room's design inherently demands more strategic and coordinated movements. If the model fails to exhibit these characteristics, it can lead to inefficiencies and frustrations, further influencing player preferences. The model's behaviour, especially in terms of movement and positioning, becomes a critical factor in this layout, potentially overshadowing other aspects of gameplay.

For a detailed visual depiction of the preference distribution in the Ring Room, readers are referred to the associated pie chart. This visualization offers deeper insights into how participants' preferences align with the scores, shedding light on their reactions to the model's performance in this particular layout.

5.3.4 Layout4: Force Coordination

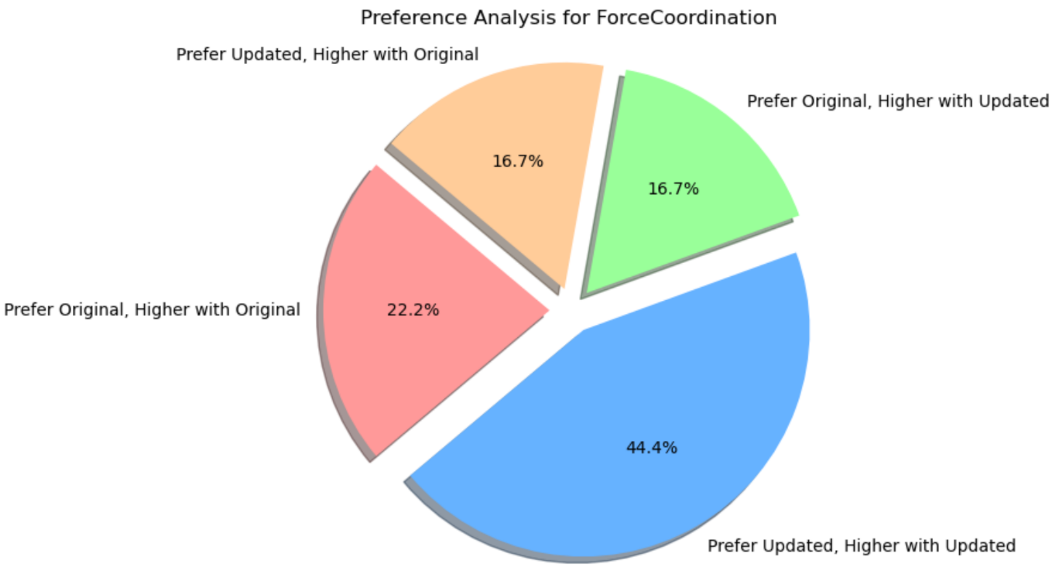

Figure 5.7: Pie Chart for Player Preferences and Score Distribution in Forced Coordination Layout

In the Forced Coordination layout, an exploration was conducted to determine the relationship between player preference and the model that achieved a higher score. The null hypothesis for this study posits that there is no significant association between player preference and the model that achieved a higher score. Conversely, the research hypothesis suggests a significant association between these two variables.

Utilising the chi-square test, a statistical value of 0.5950 was obtained. Accompanied by a p-value of 0.4405, which exceeds the standard alpha level of 0.05, the null hypothesis

Table 5.7: Correlation between Preference and Higher Score in Forced Coordination Layout

Metric/Combination	Value
Prefer original and scored higher with original	4
Prefer updated and scored higher with updated	8
Prefer original and scored higher with updated	3
Prefer updated and scored higher with original	3
Chi-square value	0.5950
P-value	0.4405

cannot be rejected. This indicates that in the Forced Coordination layout, there isn't a significant correlation between player preference and the model that achieved a higher score.

The Forced Coordination layout is characterised by a design where the participant and the model are segregated into two distinct regions, requiring collaboration to accomplish the objective. In this layout, the model has the liberty to maneuver without being hindered by the participant, such as obstructions in pathways, which is quite similar to the Cramped Room layout. The absence of such constraints might lead to a scenario where the model's actions, though contributing to the score, may not be as perceptible or impactful to the player's overall experience. This could explain the lack of a strong correlation between score and preference.

Additionally, given the inherent need for collaboration in this layout, participants might place a higher emphasis on the model's ability to understand and respond to their actions rather than the score itself. The nuances of collaboration, such as timing, anticipation, and mutual understanding, might overshadow the raw score in determining player preference.

For a comprehensive visual representation of the preference distribution in the Forced Coordination layout, readers are directed to the accompanying pie chart. This illustration provides a deeper understanding of how participants' preferences are distributed in relation to the scores they achieved.

5.3.5 Layout5: Counter Circuit

In the Counter Circuit layout, an investigation was undertaken to discern the relationship between player preference and the model that achieved a superior score. The null hypothesis for this study posits that there is no significant association between the two variables. Conversely, the research hypothesis suggests a significant correlation between player preference and the model that secured a higher score.

Using the chi-square test, a statistical value of 4.7531 was derived. With a p-value of 0.0292, which is below the conventional alpha level of 0.05, the null hypothesis is

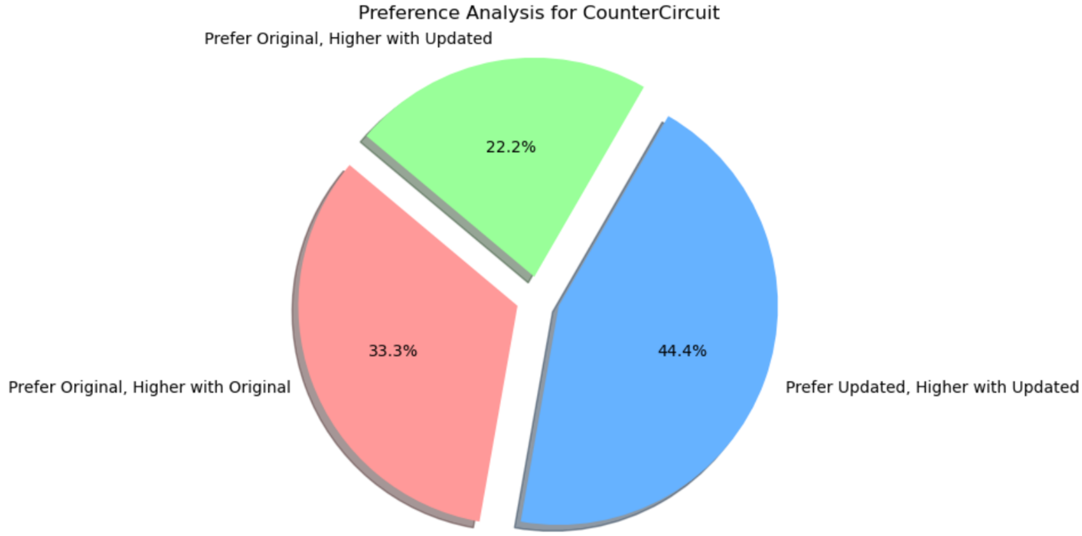

Figure 5.8: Pie Chart for Player Preferences and Score Distribution in Counter Circuit Layout

Table 5.8: Correlation between Preference and Higher Score in Counter Circuit Layout

Metric/Combination	**Value**
Prefer original and scored higher with original	6
Prefer updated and scored higher with updated	8
Prefer original and scored higher with updated	4
Prefer updated and scored higher with original	0
Chi-square value	4.7531
P-value	0.0292

rejected in favor of the research hypothesis. This indicates that in the Counter Circuit layout, there exists a significant correlation between player preference and the model that achieved a higher score.

The Counter Circuit layout is an evolved version of the Ring Room. In this layout, the model's actions, particularly if it obstructs the player's path, can lead to diminished scores. Such obstructions not only impact the score but can also significantly affect the player's overall experience. This direct influence of the model's actions on the gameplay might be a primary reason for the observed correlation between score and preference.

Furthermore, given the intricate nature of the Counter Circuit layout, players might be more attuned to the model's actions and their direct consequences on the game's outcome. If the model's actions align with the player's strategy and lead to a higher score, it could bolster the player's preference for the updated model. Conversely, any misalignment, especially if it results in a lower score, could sway the player's preference towards the original model.

5.4 Participants' Feedback

The feedback provided by the participants offers a deeper understanding of the player experience and their perception of AI teammates. Analysing this feedback in light of the experiment results can provide valuable insights into the dynamics of human-AI collaboration in gaming scenarios.

- **Perception of AI in Tense Scenarios:** The participants highlight that in high-pressure or tense game scenarios, participants often cannot discern the differences between the models. This observation aligns with the results from our experiment where, in certain layouts, there was significant correlation between preference and higher score. This suggests that in intense game situations, players might prioritise the game's outcome over the nuances of AI behaviour.
- **AI's Role in Independent Tasks:** The participants highlight the importance of AI non-interference when players can complete tasks independently. This resonates with the results from the Asymmetric Room layout, where the room did not necessarily require collaboration. Players preferred the AI that did not obstruct their strategies, even if it meant the AI being passive.
- **AI as "Assistants":** Some participants mention that it is difficult to consider the AI teammate as teammate. They expressed a preference for AI agents that complement their actions by performing tasks that the players haven't addressed. This suggests that players value AI that can fill in gaps or handle overlooked tasks, rather than AI that takes initiative or makes decisions that might disrupt the player's strategy. This perspective aligns with the results from the Ring Room and Counter Circuit layouts, where the AI's actions that are not aligning with the participants' decisions, could significantly affect their experience.

Chapter 6

Discussion

The aim of this research is to explore the potential benefits of modifying AI models that were initially trained without human data by adding adjustment layer based on human feedback when outputting the action to better align with player preferences. To address this, we incorporated human preference data and conducted experiments to gauge player inclinations. Our approach began with understanding the key elements that enhance player experience in multiplayer games Emmerich and Masuch (2016). We then developed methods to gather and interpret subjective feedback from participants, drawing inspiration from established strategies IJsselsteijn et al. (2013). For our experiments, we selected the cutting-edge Overcooked AI, PECAN, as our baseline model for enhancements and comparisons Lou et al. (2023). Participants were given the opportunity to customize their updated model and then play the game with both the original and modified versions, offering feedback on each. Through a mix of quantitative and qualitative data collection, we subsequently assessed the performance of our modified model in comparison to the original.

From the data gathered, we observed varied outcomes concerning the model updated with human feedback. When examining player scores, there wasn't a substantial difference between the outcomes achieved with the original and the updated models. This is highlighted by the average scores across all games, with the original model scoring an average of 201.20 and the updated model averaging 201.80. Further statistical analysis, including t-tests and p-value calculations, supports this observation. A p-value of 0.4743 suggests that the null hypothesis can be rejected, implying that the score differences between the two models are statistically insignificant. This aligns with our design intentions, as we made only minor adjustments to the output policy. We merely added a normalized vector with slight additional weight to certain actions, ensuring decisions would only be slightly influenced when initial probability differences were minimal. Given the advanced nature of the PECAN model and our primary goal of enhancing player preference over performance, maintaining or slightly improving the score is a positive outcome for our investigation.

The experiment revealed a slight preference among players for the updated model over the original, though the difference wasn't substantial. The cramped room saw the highest preference for the updated model, likely because participants based their adjustments on their observations in this room. Conversely, the asymmetric room, with its unique design, garnered the least preference for the updated model. This room's design, which divides the player and the model into separate areas, diminishes the need for collaboration. This shift transforms the game from a collaborative experience to a multiplayer setting that doesn't necessitate teamwork. As highlighted in Wang et al. (2021), players generally exhibit limited trust in AI teammates. In scenarios where collaboration is minimal, this inherent distrust further diminishes the AI's perceived presence. Consequently, players tend to base their preferences on their performance and the scores achieved with each model.

The observed correlation between participant preferences and scores was unanticipated. Our initial expectation was for participants to base their preferences solely on their gameplay experience. Yet, the analysis revealed a chi-square value of 24.68, suggesting a robust correlation between scores and player preferences. Feedback from participants highlighted that during intense gameplay, they often found it challenging to monitor their AI teammate's actions or discern any unexpected behaviours. While not explicitly mentioned in the feedback, literature such as Emmerich and Masuch (2016) and Norman (2013) suggests that a fulfilling game experience is closely tied to a sense of competence. In multiplayer games, this feeling often stems from contributing to the team's success. Given that "Overcooked" evaluates success based on scores, it's natural for players to derive their sense of competence from the scores they achieve. Consequently, it's understandable that players would gravitate towards the model that aids them in securing higher scores.

Examining individual layouts, a strong correlation emerges in situations where unsuitable actions have a pronounced impact. Specifically, the two layouts featuring circuits, namely the Ring room and the Counter Circuit room, both displayed a notable correlation between score and preference. However, this correlation wasn't as pronounced as in the Asymmetric room, where preference seemed entirely tied to scores. A plausible explanation for this correlation in the circuit layouts is the model's obstructive actions. When the model acts in a way that impedes the human player, it adversely affects both the progression of the task and the player's disposition. Consequently, not only does the score suffer, but the model also fails to gain the player's preference.

One of the main limitations of this experiment is that the model cannot be adjusted for each different action based on the preferences of the participants. Since the action space is designed to only consist of moving in different directions, staying still, and conducting special actions, it can only adapt feedback that is not too specific or too vague, such as "I don't like the model to do the food preparation task if possible", and "I prefer it to deliver the food". However, for feedbacks like "I don't want the model to do what I am doing", or "I want it to go in clockwise direction" are not really feasible. Unfortunately, due to the nature of reinforcement learning models, defining and training the model with all these specific actions in the action space is not efficient, which means that the adjustment to the player's output is also limited.

Another challenge faced in this study is the difficulty in generalizing feedback across different rooms, as players might have varying preferences depending on the layout. For example, if a player expresses a dislike for the model's aimless movement in one room, we might increase the weight for the "standing still" action during adjustments. However, in rooms with circuit layouts, having the model stand still in a spot, even if it doesn't significantly affect overall performance, can be perceived as obstructive. While the path isn't entirely blocked and players could opt for an alternate route, the pressure of the game can hinder on-the-fly decision-making IJsselsteijn et al. (2013). This can lead players to stick to their initial choices and perceive the AI as a hindrance, potentially affecting their overall preference negatively.

In summary, the introduced adjustment layer has shown promise, garnering higher player preferences compared to the original model. While there were unforeseen circumstances, such as players favouring models that achieved higher scores, these outcomes can be rationalized based on prior literature reviews. One potential solution to reduce score influence on player preference could be to hide the scores, but considering the game's design where success is directly tied to scores, displaying them aligns with the intended game experience. The study faced challenges, particularly in the precise adjustment of the model due to the inherent characteristics of reinforcement learning agents and their action space design. Additionally, feedback seemed to be scenario-specific, limiting its broader applicability. Given that our tests were confined to five scenarios, it's premature to draw definitive conclusions. For a more comprehensive understanding of the model's adaptability, future research should explore a wider range of game scenarios with varied layouts and tasks.

Chapter 7

Conclusion and Future Work

7.1 Conclusion

The primary objective of this research was to investigate the impact of incorporating human feedback for model adjustments on player preferences in a collaborative game environment. Drawing from the literature review, we synthesised key insights from prior studies into actionable guidelines for AI model design and evaluation in collaborative settings. Using the original PECAN model as our foundation, the participants reviewed and analysed its performance and provided feedback, leading to the development of an updated model with specific adjustments based on player feedback.

Our evaluations demonstrated that the updated model obtained a higher preference among players and achieved a slightly better average score than the original. Quantitative data revealed a correlation between score and preference, suggesting that players tend to favour models that yield higher scores. This not only underscores the effectiveness of the updated model but also highlights its enhanced collaboration with players.

However, while the updated model showed promising results, it is not without areas for improvement. Feedback from participants during the experiment indicated that in certain game scenarios, players couldn't discern significant differences between the models, and this became one of the main reasons for players opted for the model with a higher score. This score, however, is also relevant to the player's performance in each game trial. Since we cannot confirm that the player performance is consistent throughout the experiment with each model, it is difficult to conclude on whether the higher score is obtained due to the better performance or collaboration performed by the model, or just the operation fluctuation caused by the human player.

In conclusion, this research provides valuable insights into the dynamics of player-AI collaboration in-game environments. While the updated model exhibits improved collaboration and player preference, there remains potential for further refinement. Nevertheless, this work serves as a foundational step for future endeavours aiming to optimise AI models in collaborative gaming contexts, ensuring both performance and player satisfaction.

7.2 Future Work

Building on the findings of this research, there are several potential paths to further explore from the current findings. Below are the future works that we suggest:

- Instead of a single session evaluation, participants could have an extended trial period where they can interact with both AI models over multiple sessions. This extended interaction might yield deeper insights into player preferences and the AI model's performance over time, and reduce the effects caused by the inconsistency of human player's performance.
- Replicating the methodology used in this research on other game layouts or entirely different game genres. This would help in understanding if the proposed method of using the human data as an adjustment factor after training the model is specific to the current game scenarios or is more universally applicable.
- Create and test the model in more scenarios to validate the generalizability of the adjustment made to the model in various scenarios.

Appendix A

Appendix: Experiment Data

A.1 Score Table with the Original Model

	Participants	CrampedRoom	AsymmetricRoom	RingRoom	ForceCoordination	CounterCircuit
1	1	200	260	140	160	140
2	2	220	280	140	160	140
3	3	200	300	160	180	140
4	4	220	320	160	180	160
5	5	180	320	180	160	140
6	6	260	300	160	160	160
7	7	260	280	140	200	180
8	8	280	260	160	220	160
9	9	300	300	180	220	140
10	10	260	320	180	140	180
11	11	240	260	160	160	160
12	12	220	360	140	140	120
13	13	200	380	120	160	160
14	14	200	300	180	180	140
15	15	240	360	200	120	160
16	16	220	300	100	180	100
17	17	240	320	140	140	140
18	18	240	340	120	160	120
19	19	240	300	160	140	160
20	20	200	320	140	160	140

A.2 Score Table with the Updated Model

	Participants	CrampedRoom	AsymmetricRoom	RingRoom	ForceCoordination	CounterCircuit
1	1	220	220	160	180	160
2	2	240	260	160	140	160
3	3	220	280	120	200	180
4	4	240	280	140	200	140
5	5	180	280	160	140	120
6	6	280	260	180	140	180
7	7	280	260	160	180	140
8	8	280	280	180	240	180
9	9	300	320	120	240	120
10	10	280	260	200	240	200
11	11	280	240	140	180	180
12	12	240	380	160	120	140
13	13	200	340	100	180	140
14	14	200	280	200	200	140
15	15	220	320	200	140	140
16	16	200	320	160	140	180
17	17	220	300	120	140	140
18	18	220	320	120	180	140
19	19	220	320	140	140	180
20	20	180	340	120	140	160

A.3 Player preference Table

Participants	CrampedRoom	AsymmetricRoom	RingRoom	ForceCoordination	CounterCircuit
1	1	0	1	1	1
2	1	0	1	0	1
3	1	0	0	1	1
4	0	0	0	1	0
5	1	0	0	0	0
6	0	0	1	1	1
7	1	0	1	0	0
8	0	1	1	0	1
9	1	1	0	1	0
10	1	0	1	1	1
11	1	0	0	1	1
12	0	1	1	0	1
13	0	0	0	1	0
14	1	0	1	0	1
15	0	0	0	1	0
16	1	1	1	1	0
17	1	0	1	1	0
18	0	0	0	0	0
19	1	1	1	0	0
20	1	1	1	1	0

Appendix B

Appendix: Experiment Materials

HUMAN RESEARCH ETHICS COMMITTEE
Application Form

Created by: **u7227895**
Record number: **14546**
Protocol type: **Expedited Ethical Review (E1)**
Protocol number: **H/2023/1195**

Date entered: **21/07/2023**
Ethics program type: **Undergraduate**
Requested start date: **19/09/2023**
Requested end date: **10/10/2023**

Protocol title: **Enhancing Collaboration in Overcooked: Improving AI-Agent Adaptability to Human Player Preferences**

Investigators

Name	Role	Department
Kyburz, Penny	Supervisor	CECC School of Computing, CECC School of Computing, ANU
Liu, Yuechen	Primary investigator	CECC School of Computing, CECC School of Computing, ANU

Investigators Detailed

Name: Liu, Yuechen **Role:** Primary investigator

Expertise: I am a third-year bachelor student in the School of Computing at ANU. During my bachelor degree, I have completed COMP2550 course Advanced Research Method and obtained a High Distinction Grade. Through completing this course, I have learned multiple techniques for conducting research, as well as skills and knowledges for data collection. I believe that the knowledge and experience gained through the course could provide sufficient support for the current methodologies I have decided to conduct for my research project.

Name: Kyburz, Penny **Role:** Supervisor

Expertise: Dr Kyburz has used methods involving qualitative, quantitative, and biometric data collection and analysis related to video games and HCI, including laboratory studies and large-scale online surveys. She has led usability and playability testing for major video game publishers in Australia and the US. She has served as an Associate Chair for CHI, the premier international conference on HCI, as well as CHI PLAY (CHI for games). She is a Technical Program Co-Chair for Designing Interactive Systems 2022 and General Co-Chair for OzCHI

2022. She is an Associated Editor for the journals Entertainment Computing and IEEE Transactions on Games and has served as a program committee member and reviewer for dozens of conferences, journals, and books. She is a former board member of Digital Rights Watch, which is an advocacy organisation for human rights in the digital space.

External Investigators

Name	Role	Institution

Departments

Primary	Department	Faculty
Yes	CECC School of Computing	CECC School of Computing

Project Questions Detailed

Description of Project

Describe the research project in terms easily understood by a lay reader, using simple and non-technical language. The purpose of this research project is to enhance the performance of the AI game players to collaborate better with human players according to their preferences. In order to evaluate the performance of these AI players, a simulator of the collaborative video game called Overcooked will be used. In the simulator, two players will work collaboratively in a small kitchen to prepare food and to serve the customer as quick as possible.

The current state-of-the-art techniques for training these AI game players gets them to play the game simulator repetitively with themselves or other AI agents as teammates, and make modifications to their decision to increase the score of each game play. Although the ultimate goal is to create AI agents that cooperate with human players, human data are not used in the training process since the collection is expensive. Because of this, although these agents can achieve satisfactory performance in terms of scores and task achievements after training, they struggle to collaborate with human players in a manner that aligns with the players' preferences.

This research project focuses on enhancing the collaborative capabilities of an AI agent within the game "Overcooked." More specifically, the aim is to develop an agent that not only understands the game's mechanics but also aligns its actions with the preferences of human players. To achieve this, the decision-making stage of the trained model for generating the action will be adjusted so that it tends to react in the way that the players prefer. The ultimate goal is to create an AI teammate that can adapt to and resonate with the unique playing preferences of its human teammates, fostering a more harmonious gameplay experience.

HUMAN RESEARCH ETHICS COMMITTEE
Application Form

Location of Data Collection

Australia Yes

Overseas No

Provide country / area where data collection will be conducted The Australian National University, Canberra, ACT 0200

Aims of the Project

List the hypothesis and objectives of your research project. The hypothesis of this research project is that by incorporating human player preferences into the decision-making process of the trained model, the Overcooked AI agent will exhibit improved adaptability and receive a higher preference rate from human players than the same agent but without involving human data for decision making. We believe that taking the insights from human players into consideration will enhance the gameplay experiences of the human players.

The objective of this research project is to develop an AI teammate for "Overcooked" based on the state-of-the-art agents. The updated model will factor in human player preferences, and the measure of success will be the preference shown by actual players towards this updated AI teammate.

Methodology

In language appropriate for a lay reader, explain why the methodological approach minimises the risk to participants. (For surveys, include justification of the sample size).

The research begins with a data collection gathered from human players about their preferences when collaborating with AI teammates in a gaming environment. This feedback will then be used for the refinement of the existing AI model, such that it can better align with human expectations. The refined model is then tested in various game scenarios, with participants interacting with both the original and updated AI versions without knowing the order of the agents. After playing with both agents in each scenario, the participants will be asked to compare the two agents and indicate a preference between the two agents. After finishing all the games, the participants will be asked to reflect on their overall gameplay experience, and provide additional feedback and suggestions on how the AI agents in "Overcooked" can be improved to enhance the overall gameplay experience.

In this research project, we will ensure the well-being and comfort of participants while playing the simulator collaboratively with the AI agents. Since Overcooked is rated as PG and does not involve mature themes, the risk to participants is significantly reduced. To mitigate any potential discomfort, I will offer participants a comprehensive overview of the experiment before its commencement, detailing the procedures and expectations. Participants will also be informed that they have the autonomy to pause or terminate the experiment at any time. Through providing sufficient information to the participants and granting them the flexibility to pause or withdraw, I aim to create a safe and respectful research environment while gathering valuable feedbacks from the participants on how this approach can effectively improve AI agents' adaptability to human player preferences, and enhance the cooperative gameplay experiences.

Provide the survey method, a list of the questions to be asked or an indicative sample of questions. These should give a good sense of the most intrusive/sensitive areas of questioning. Please find a copy of the indicative experiment flow for the participants attached in the documents. The document includes the estimated experiment duration, instructions to the participants, and sample questions for feedback after the experiment.

What mechanisms do the researchers intend to implement to monitor the conduct and progress of the research project? For example:
How often will the researcher be in touch with the supervisor?
Is data collection going as expected? If not, what will the researcher do?
Is the recruitment process effective?
How will the researcher monitor participants' willingness to continue participation in the research project, particularly when the research is ongoing?
I plan to conduct the research experiments over a three-week period, and at the end of each week, I will provide an update to my supervisor regarding the progress of data collection. During this time, I will keep track of the number of experiments conducted and the remaining ones. If I notice that data collection is not progressing as expected, I will promptly inform my supervisor and discuss potential options for proceeding, which may involve organizing additional experiments to gather more data if necessary.
By maintaining regular communication with my supervisor, closely monitoring data collection progress, and establishing trust with participants through effective recruitment, I aim to conduct a successful research project that respects participants' comfort and contributes valuable insights to the study's objectives.
In order to monitor the willingness of the participants to continue their participation in the research project, I will notify the participants through the information sheet, and orally before the experiment that they will be allowed to quit the participation freely at any time before the submission of the assessment. During the experiment, I will ensure that the participants can communicate with me about any concerns at any time. I will also observe the condition of the participants to see if they show any signs of discomfort, and ask them for their willingness to continue the experiment after each trial of comparisons.

Participants

Provide details in relation to the potential participant pool, including:

target participant group;
identification of potential participants;
initial contact method, and
recruitment method. The target participant group for this research project includes individuals who have experience playing the video game Overcooked. These players should be familiar with the game mechanics and have a reasonable understanding of the cooperative nature of gameplay. We welcome adults from any professional background to participate. I will display posters around ANU detailing the target participant group's information. These posters will also provide my phone number and email address, allowing potential participants to reach out and inquire about any specific details regarding the research.

Proposed number of participants 20

Provide details as to why these participants have been chosen? The participants for this research project have been carefully chosen based on their relevant experience and familiarity

HUMAN RESEARCH ETHICS COMMITTEE
Application Form

with the video game Overcooked. By targeting individuals who have played the game, we can benefit from their insights and opinions related to the game mechanics and the cooperative nature of gameplay. Their firsthand experiences in Overcooked will allow them to provide informed feedback on AI teammates and their preferences for collaborative gameplay.

Potential participants include individuals from various backgrounds, disciplines, and professions. Engaging participants from different walks of life ensures a comprehensive understanding of the research topic, incorporating viewpoints from individuals with different gaming habits and preferences. This inclusive strategy aims to capture a wide range of perspectives, enriching the research findings with diverse insights.

Cultural and Social Considerations/Sensitivities

What cultural and/or social considerations/sensitivities are relevant to the participants in this research project? There won't be any cultural or social sensitivities relevant to the participants in this research project.

Incentives

Will participants be paid or any incentives offered? If so, provide justification and details. It is unlikely that the participants will be paid or offered with any incentives from participating in this research. However, the work done in this research may contribute to the advancement of AI algorithms, not only within Overcooked but also in other games and applications. The insights gained from player feedback and preferences will inform smarter AI design and implementation, impacting AI technologies in real-world scenarios beyond gaming.

Benefits

What are the anticipated benefits of the research? The primary benefit of this research is that it could provide an alternative way to train the AI game agents to adapt to human players more without totally relying on human data. This would avoid the expensive cost for data collection, as the AI agent will still be trained without human data at the stage for understanding the rules and mechanisms of the game, and only use the human data to change the final decision-making process such that the agents tend to select the action that is favoured by human players from several appropriate actions for the game state. Consequently, the players will enjoy enhanced gaming experiences as AI agents adapt to their preferences more, creating a more personalised and engaging gameplay.

To whom will the benefits flow? Players will be the primary beneficiaries of the research's outcomes. They will experience enhanced gaming experiences with AI agents that adapt to their preferences, resulting in more enjoyable and engaging gameplay.
The research's insights and recommendations will also benefit game developers. By taking the human data into consideration for the final decision making stage of AI agents that are already trained to understand the rules and mechanisms, developers can create more personalised and adaptive AIs in games without significantly increasing the workload.

Informed Consent

Indicate how informed consent will be obtained from participants. At least one of the

following boxes MUST be ticked 'Yes'.

In writing Yes

Return of survey or questionnaire No

Orally No

Other No

If Oral Consent or Other, provide details.

Confidentiality
Confidentiality will be protected as far as the law allows.

Describe the procedures that will be adopted to ensure confidentiality during the collection phase and in the publication of results. During the data collection phase, participants' personal information will be anonymised and stored securely to ensure confidentiality. Access to the data will be restricted to authorized researchers only (me and my supervisor Dr. Penny Kyburz). In the publication of results, all data will be presented in aggregate form, preventing the identification of individual participants. Pseudonyms may be used to further protect participant identities and maintain confidentiality.

Data Storage Procedures

Provide an overview of the data storage procedures for the research. Include security measures and duration of storage. All digital data including game trajectories and the responses of the participants will be securely stored on my ANU OneDrive, ensuring limited access only to myself and my supervisor. No other individuals will be granted permission to access or edit the collected data.
Given that the data will be used for an undergraduate project, it will be stored by the primary investigator for one year in order to comply with university policy. At the end of this period, all digital data held by the primary investigator will be de-identified and archived with the permission of the participants. Otherwise, the data will be permanently destroyed.
Feedback

Provide details of how the results of the research will be reported / disseminated, including the appropriate provision of results to participants. If appropriate, provide details of any planned debriefing of participants. The data collected from this research will serve as the foundation for a thesis and a concise presentation, shared with other students and academics. Additionally, the findings may be expanded into a research paper if possible. Survey responses will be reported in aggregate form to provide overall insights, and individual participant feedback will be kept confidential and non-identifiable to respect their privacy and confidentiality.

Supporting Documentation

Have you uploaded all relevant supporting documentation, such as Participant Information Sheet and/or consent form, to the documents tab?
Yes

HUMAN RESEARCH ETHICS COMMITTEE
Application Form

Has this work been approved by another Human Research Ethics Committee (HREC)? No

If yes, please give the name of the approving HREC. You will also need to include a copy of the approval letter in your application and also upload an electronic copy to the Documents tab.

Funding

Is this research supported by external funding? No

Provide the name/s of the external sources of funding. Please include grant number/s if available.

Is the research conducted under the terms of a contract of consultancy agreement between the ANU and the funding source? No

Describe all the contractual rights of the funding source that relate to the ethical consideration of the research.

HUMAN RESEARCH ETHICS COMMITTEE
Application Form

Expedited Questions Summary

Question	Answer
Third Party Identification	No
Children or Young People	No
Dependent or Unequal Relationship	No
Membership of a Group, or Related Issues	No
Physical Harm	No
Psychological Harm (includes Devaluation of Personal Worth)	No
Social Harm	No
Economic Harm	No
Legal Harm	No
Covert Observation	No
Deception	No
Sensitive Personal Information	No
Overseas Research	No
Secondary Data	No
Collection, use or disclosure of personal information WITHOUT the consent of the participant	No

HUMAN RESEARCH ETHICS COMMITTEE
Application Form

Supporting Documentation

Please ensure electronic copies of the supporting documentation have been uploaded into the documents tab of your protocol.

These may include (please circle the relevant answer):

List of indicative questions Y/N

Copy of questionnaire / survey Y/N

Invitation or introductory letter/s Y/N

Publicity material (posters etc.) Y/N

Information sheet Y/N

Consent form Y/N

External approval documentation Y/N

Research visa (if applicable) Y/N

Other (specify below) Y/N

For other, please specify:

Page **9** of **10**

HUMAN RESEARCH ETHICS COMMITTEE
Application Form

SIGNATURES AND UNDERTAKINGS

PROPOSER OF THE RESEARCH

I certify that all the persons listed in this protocol have been fully briefed on appropriate procedures and in particular that they have read and are familiar with the national guidelines issued by the National Health and Medical Research Council (the National Statement on Ethical Conduct in Human Research 2007).

I certify that the above is as accurate a description of my research proposal as possible and that the research will be conducted in accordance with the National Statement on Ethical Conduct in Human Research 2007. I also agree to adhere to the conditions of approval stipulated by the ANU Human Research Ethics Committee (HREC) and will cooperate with HREC monitoring requirements. I agree to notify the Committee in writing immediately of any significant departures from this protocol and will not continue the research if ethical approval is withdrawn and will comply with any special conditions required by the HREC.

Signed:.. Date:..................

ANU SUPERVISOR

I certify that I shall provide appropriate supervision to the student to ensure that the project is undertaken in accordance with the undertakings above:

Signed:.. Date:..................

AS FROM MONDAY 21ST OCTOBER 2013 THE SIGNATURE OF THE HEAD OF ANU DEPARTMENT/GROUP/CENTRE IS NO LONGER REQUIRED.

Australian National University

Participant Information Sheet

Researcher:

My name is Yuechen Liu. I am a third-year bachelor student in the School of Computing at ANU, with my specialization being Machine Learning.

Project Title: Enhancing Collaboration in Overcooked: Improving AI-Agent Adaptability to Human Player Preferences

General Outline of the Project:

- **Description and Methodology:** The purpose of this research project is to enhance the performance of artificial intelligence (AI) game players to collaborate effectively with human players according to their preferences. The original AI agent only knows the rules and possible actions of the game, but have no idea about what human players prefer. In this research, the original AI agent will be adjusted using the information collected from human players regarding their teammate preferences. Consequently, the agent is expected to choose the most suitable action based on what the players prefer from a set of possible actions at a given game state.

  The research begins with a data collection gathered from human players about their preferences when collaborating with AI teammates in a gaming environment. This feedback will then be used for the refinement of the existing AI agent, such that it can better align with human expectations. The refined agent is then tested in various game scenarios, with participants interacting with both the original and updated versions. Their feedback determines the effectiveness of the modifications and provides direction for any further enhancements, ensuring the AI resonates with human player preferences and expectations.

- **Participants:** 20 participants will be invited to play two versions of the simulator of the video game Overcooked, and provide feedback. The target participant group for this research project includes individuals who have experience playing the video game Overcooked. These players should be familiar with the game mechanics and have a reasonable understanding of the cooperative nature of gameplay. The potential participant will be identified and recruited, and their ages should be 18 years or older. There will be no limitations in the disciplines or professions of the participants.

- **Use of Data and Feedback:** The data collected from this research will be used to produce the final evaluation section of both a project report and a short presentation. After the completion of the research project, the participants can contact me via email to access the outcome of the research project.

Participant Involvement:

- **Voluntary Participation & Withdrawal:**
  Participation in this study is entirely voluntary, and the participants have the freedom to decline or withdraw from the experiment without needing to provide any explanation, up until the work is prepared for the submission of the assessment. If the participants decide to withdraw, they may do so within a

period of 2 weeks after the experiment. In such a case, any data provided prior to withdrawal will be destroyed and will not be used in the research.

- **What does participation in the research entail?**
  The process begins with a demonstration that illustrates the AI's collaborative actions in the game presented by the agent that the participants will interact with. After observing, participants will provide feedback on the actions of the agents, highlighting preferred behaviours and areas for improvement. Based on this feedback, the primary researcher will adjust the AI to better align with participants' preferences.

  Subsequently, participants will engage with two versions of the agents in a game simulator without knowing what exactly they are. This engagement encompasses five distinct scenarios, each offering a unique collaborative experience with an AI teammate. After these interactions, participants will evaluate the AI's performance, indicating a preference and offering insights on their overall gameplay experience.

- **Location and Duration:** Experiments will be conducted at the CSIT building, and are expected to last for 40 to 60 minutes.

- **Risks:** The research involves gathering feedback and preferences from human players regarding their experiences with AI agents. Discussions or interactions related to gameplay and AI behaviour could evoke emotional stress or frustration. To address the potential risks and discomforts, informed consent will be obtained, and the purpose, procedures, and any potential risks associated with the experiment will be explained to the participants prior to the experiment. Participants will also have the option to withdraw from the study at any time.

- **Benefits:** It is unlikely you will derive any substantial personal benefit from participation in this research. However, the work done in this research may contribute to the advancement of AI algorithms, not only within Overcooked but also in other games and applications. The insights gained from player feedback and preferences will inform smarter AI design and implementation, impacting AI technologies in real-world scenarios beyond gaming.

Confidentiality:

- **Confidentiality:** During the data collection phase, participants' personal information will be anonymized and stored securely to ensure confidentiality. Access to the data will be restricted to authorised researchers only (me and my supervisor Dr. Penny Kyburz). In the publication of results, all data will be presented in aggregate form, preventing the identification of individual participants. Pseudonyms may be used to further protect participant identities and maintain confidentiality.

Privacy Notice:

In collecting your personal information within this research, the ANU must comply with the Privacy Act 1988. The ANU Privacy Policy is available at https://policies.anu.edu.au/ppl/document/ANUP_010007 and it contains information about how a person can:

- Access or seek correction to their personal information;
- Complain about a breach of an Australian Privacy Principle by ANU, and how ANU will handle the complaint.

Australian National University

Data Storage:

- **Where:** All digital data and audio recordings from interviews will be securely stored on ANU OneDrive, ensuring limited access only to myself and my supervisor. No other individuals will be granted permission to access or edit the collected data.

- **How long:** Given that the data will be used for undergraduate research, all research data will be retained and stored for one year following the submission of the research.

- **Handling of Data following the required storage period:** At the end of this period all digital data held by the primary investigator will be de-identified and archived with the permission of the participants. Otherwise, the data will be permanently destroyed.

Queries and Concerns:

- **Contact Details for More Information:** Please contact the primary investigator Yuechen Liu via u7227895@anu.edu.au (+61 468491061), or the supervisor Dr Penny Kyburz via penny.kyburz@anu.edu.au (+61 2 6125 1607) if there are any further queries regarding the project.

Ethics Committee Clearance:

The ethical aspects of this research have been approved by the ANU Human Research Ethics Committee (Protocol H/2023/1195). If you have any concerns or complaints about how this research has been conducted, please contact:

Ethics Manager
The ANU Human Research Ethics Committee
The Australian National University
Telephone: +61 2 6125 3427
Email: Human.Ethics.Officer@anu.edu.au

Indicative Interview & Experiment Guide

Overall, this research session is designed to last between 40 to 60 minutes, where the participants are expected to interact with the collaborative AI agents to experience the overcooked game in 5 different scenarios. The experiment will start by presenting a demonstration to the participants showcasing how the AI behaves when they are collaborating with itself in games. In this stage, the participants are encouraged to watch this demonstration closely and imagine these AI agents as their daily teammates. After the demonstration, they will be asked about the specific actions they'd like to see from the AI teammates and any behaviours they'd prefer the AI not to exhibit.

Hence, the agent will be slightly modified to adapt to the player's preferences. Afterwards, the participants will interact with two AI models: the original model they have observed, and the updated model, within a game simulator. There will be five different scenarios for the participants to navigate through. The participant will team up with each model once in each scenario, so they will participate in 10 trials in total. The trials will be counterbalanced such that the participants will be presented with the AIs in different and unspecified order for each trial.

After interacting with the two AIs in each scenario, participants are prompted to provide feedback on the performances of the agents, and indicate a preference between the two agents. After finishing all the games, the participants will be asked to reflect on their overall gameplay experience, and provide additional feedback and suggestions on how the AI agents in "Overcooked" can be improved to enhance the overall gameplay experience.

Here is an indicative interview and experiment guide.

<u>Greeting and Introduction (3 min):</u>

Thank you for participating in my research on enhancing collaboration between AI and human players in Overcooked. Here is the consent form for you to agree with participating in this experiment. During this experiment, your game scores and your responses will be saved with your consent, and you'll have the opportunity to review them afterward. If any question in the interview or phases in games make you uncomfortable, feel free to skip it or pause the experiment. Take your time to respond, and don't hesitate to ask for clarification if needed. Your well-being is essential, and if you need a break, please let me know. If you don't have any questions, please sign the consent form.

Thank you again for your participation, and let's begin the interview.

<u>Teammate preference setup (5 minutes):</u>

I will firstly present you with a short demonstration on how the AI you will be playing with interact with another Overcooked AI. Please carefully watch the demonstration and imagine the AI as your Overcooked teammates, and look for the actions you don't want them to exhibit. You are also allowed to pause or restart the demonstration, and discuss with me at any time about the actions of taken by the AI if you want. If you don't have any questions, I will start playing the demonstration to you now.

<u>Experiment (5 - 8 min for each scenario, around 30 minutes in total)</u>

Thanks for your answers. Now we shall start the experiment. In this stage, you will interact with two AIs respectively in the five scenarios using a game simulator. Each trial will take 100 seconds. In order to make it fair for you the compare the two AIs, I am not going to tell you which AI comes first for each scenario. You will be the green character, and your AI teammates will be the blue one. You can control the movement of your character using the four arrow keys, and perform any special actions such as picking up the plates etc using the space key. Your final score of each trial will be saved. If you have any questions or feel uncomfortable with the experiment, feel free to stop at any time. Now let's begin with the first scenario. I have set up the game simulator for you with one of the AIs as your teammate, and you can click start once you are ready.

(Participants interacting with one of the AIs without knowing whether it is the original version or adjusted version)

Is there anything you would like to comment on for your collaboration experience with this agent?

Now we shall interact with the other agent.

(Participants interacting with the other AI without knowing whether it is the original version or adjusted version)

Is there anything you would like to comment on for your collaboration experience with this agent?

Just focus on the two experiment trials, which agent of the two do you prefer more? Do you have a reason for your preference?

Now let's move onto the next scenario. (Five in total)

Closing (3 min)

Thanks for your participation and here is our last question for today's experiment, is there anything else you would like to add or any suggestions you have for improving the AI agents in Overcooked to enhance the overall gameplay experience?

Debriefing (3 min)

Thank you once again for sharing your experiences and insights during this interview. Your valuable input will greatly contribute to the success of this research project. If you have any further thoughts or questions after the interview, please feel free to reach out. Thank you for being a part of this study, and your contribution will undoubtedly make a significant impact.

Australian
National
University

CONSENT FORM for research participants

Enhancing Collaboration in Overcooked: Improving AI-Agent Adaptability to Human Player Preferences

I have read and understood the Information Sheet you have given me about the research project, and I have had any questions and concerns about the project (listed here

____

____)

addressed to my satisfaction.

I agree to participate in the project. YES ☐ NO ☐

I agree to my response and game scores in this experiment being recorded and securely saved on ANU OneDrive.

YES ☐ NO ☐

I agree to the information I provide in this experiment being used in an undergraduate research project, with the potential for It to be developed into a research paper and published later.

YES ☐ NO ☐

Signature:……………………………………………. Date:…………………………………………………

Bibliography

Adachi, P. J. C. and Hodson, G., 2018. Playing well with others: The role of opponent and intergroup anxiety in the reduction of prejudice through collaborative video game play. (2018). https://psycnet.apa.org/record/2018-48744-001. [Cited on page 10.]

Ashktorab, Z.; Liao, Q. V.; Dugan, C.; Johnson, J.; Pan, Q.; Zhang, W.; Kumaravel, S.; and Campbell, M., 2020a. Human-ai collaboration in a cooperative game setting: Measuring social perception and outcomes. *Proc. ACM Hum.-Comput. Interact.*, 4, CSCW2 (oct 2020). doi:10.1145/3415167. https://doi.org/10.1145/3415167. [Cited on page 6.]

Ashktorab, Z.; Liao, V.; Dugan, C.; Johnson, J.; Pan, Q.; Zhang, W.; Kumaravel, S.; and Campbell, M., 2020b. Human-ai collaboration in a cooperative game setting: Measuring social perception and outcomes. *Proceedings of the ACM on Human-Computer Interaction*, 4 (10 2020), 1–20. doi:10.1145/3415167. [Cited on page 13.]

Carroll, M.; Shah, R.; Ho, M. K.; Griffiths, T. L.; Seshia, S. A.; Abbeel, P.; and Dragan, A., 2020. On the utility of learning about humans for human-ai coordination. [Cited on page 7.]

Csikszentmihalyi, M., 1990. *Flow: The Psychology of Optimal Experience.* [Cited on page 12.]

Emmerich, K. and Masuch, M., 2016. Game metrics for evaluating social in-game behavior and interaction in multiplayer games. 1–8. doi:10.1145/3001773.3001793. [Cited on pages 11, 31, and 32.]

Fox, C. and Brockmyer, J., 2013. The development of the game engagement questionnaire: A measure of engagement in video game playing: Response to reviews. *Interacting with Computers*, 25 (06 2013), 290–293. doi:10.1093/iwc/iwt003. [Cited on page 12.]

Games, G. T., 2016. Overcooked. [Cited on page 6.]

Hopson, J., 2001. Rethinking carrots: A new method for measuring what players find most rewarding and motivating about your game. https://www.gamedeveloper.com/design/rethinking-carrots-a-new-method-for-measuring-what-players-find-most-rewarding-and-motivating-about-your-game. Accessed: [Your Access Date Here]. [Cited on page 11.]

Hoshino, Y.; Takagi, T.; Profio, U.; and Fujita, M., 2004. Behavior description and control using module for personal robot. vol. 4, 4165 – 4171 Vol.4. doi:10.1109/ROBOT.2004.1308925. [Cited on page 3.]

IJsselsteijn, W.; de Kort, Y.; and Poels, K., 2013. *The Game Experience Questionnaire*. Technische Universiteit Eindhoven. [Cited on pages 11, 12, 31, and 33.]

Johnson, D.; Gardner, M. J.; and Perry, R., 2018. Validation of two game experience scales: The player experience of need satisfaction (pens) and game experience questionnaire (geq). *International Journal of Human-Computer Studies*, 118 (2018), 38–46. doi:https://doi.org/10.1016/j.ijhcs.2018.05.003. https://www.sciencedirect.com/science/article/pii/S1071581918302337. [Cited on pages 10 and 11.]

Khadpe, P.; Krishna, R.; Fei-Fei, L.; Hancock, J. T.; and Bernstein, M. S., 2020. Conceptual metaphors impact perceptions of human-ai collaboration. *Proc. ACM Hum.-Comput. Interact.*, 4, CSCW2 (oct 2020). doi:10.1145/3415234. https://doi.org/10.1145/3415234. [Cited on page 1.]

Lou, X.; Guo, J.; Zhang, J.; Wang, J.; Huang, K.; and Du, Y., 2023. Pecan: Leveraging policy ensemble for context-aware zero-shot human-ai coordination. [Cited on pages 1, 5, 7, 8, 9, 10, and 31.]

Ltd., G. T. G., 2016. Overcooked. https://store.steampowered.com/app/448510/Overcooked/. [Cited on page 6.]

Norman, K., 2013. Geq (game engagement/experience questionnaire): A review of two papers. *Interacting with Computers*, 25 (06 2013), 278–283. doi:10.1093/iwc/iwt009. [Cited on page 32.]

Ravaja, N.; Saari, T.; Salminen, M.; Laarni, J.; and Kallinen, K., 2006. Phasic emotional reactions to video game events: A psychophysiological investigation. *Media Psychology - MEDIA PSYCHOL*, 8 (11 2006), 343–367. doi:10.1207/s1532785xmep0804_2. [Cited on page 12.]

Rosero, A.; Dinh, F.; de Visser, E. J.; Shaw, T.; and Phillips, E., 2021. Two many cooks: Understanding dynamic human-agent team communication and perception using overcooked 2. [Cited on pages 6 and 7.]

Ryan, R.; Rigby, C.; and Przybylski, A., 2006. The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion*, 30 (12 2006), 344–360. doi:10.1007/s11031-006-9051-8. [Cited on page 12.]

Sherry, J.; Greenberg, B.; Lucas, K.; and Lachlan, K., 2006. *Video game uses and gratifications as predictors of use and game preference*, vol. 8, 213–224. [Cited on page 12.]

Strouse, D.; McKee, K. R.; Botvinick, M.; Hughes, E.; and Everett, R., 2022. Collaborating with humans without human data. [Cited on pages 4 and 9.]

Voiskounsky, A.; Mitina, O.; and Avetisova, A., 2004. Playing online games: Flow experience. *PsychNology Journal*, 2 (01 2004), 259–281. [Cited on page 12.]

Wang, D.; Churchill, E.; Maes, P.; Fan, X.; Shneiderman, B.; Shi, Y.; and Wang, Q., 2020. From human-human collaboration to human-ai collaboration: Designing ai systems that can work together with people. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI EA '20 (Honolulu, HI, USA, 2020), 1–6. Association for Computing Machinery, New York, NY, USA. doi:10.1145/3334480.3381069. https://doi.org/10.1145/3334480.3381069. [Cited on pages 5 and 13.]

Wang, T.; Zeng, L.; Dong, W.; Yang, Q.; Yu, Y.; and Zhang, C., 2022. Context-aware sparse deep coordination graphs. [Cited on page 3.]

Wang, Z.; Huang, X.; Zhang, X.; and Zhu, Y., 2021. An ideal human" expectations of ai teammates in human-ai teaming. (2021). https://dl.acm.org/doi/abs/10.1145/3432945. [Cited on pages 1 and 32.]

Zhou, R.; Muise, C.; and Hu, T., 2022. A permutation-invariant representation of neural networks with neuron embeddings. https://openreview.net/forum?id=vuw072gfi3W. [Cited on pages 4 and 8.]