Utrecht University
Graduate School of Natural Sciences

Master thesis

Facing Your Self: Dynamic Difficulty Adjustment in Hades via Deep Player Behavior Modeling

Hui En Lin [8098735]
h.lin2@uu.students.nl

October 2023 - July 2024

Supervisors
dr. J. Pfau
dr. S.C.J. Bakkes

Abstract

This master's thesis explored the development and evaluation of enhancing Dynamic Difficulty Adjustment (DDA) with Deep Player Behavior Modeling (DPBM) in the context of video games. The DPBM approach of this thesis is designed to capture and replicate individual player behaviors, providing a personalized and engaging gaming experience. To test the effectiveness, I developed a mod named "Dark Zagreus" for the award-winning action roguelike game, Hades. In this modified version, the final boss is replaced with Dark Zagreus, an AI agent that leverages DPBM to learn from the player's last successful run and imitate player behavior. Within a study long lasted two weeks, 20 players ($n = 20$) participated in an approximately one-hour session, where they encountered the boss in both scripted and DPBM modes and filled out a questionnaire. At the end of the study, a fluctuating result was observed and no significant differences regarding player experience metrics and imitation scores were found. In the open-ended responses to the questionnaire, a portion of participants noted that Dark Zagreus effectively mirrored their combat tactics, suggesting successful behavior replication. Regarding the performance, DPBM shows a comparable accuracy with similar previous works, which still brings this research valuable insight. Despite limitations such as the exclusion of certain game mechanics, this master's thesis demonstrated the potential of DPBM in advancing DDA research. Aside from this contribution, the source code of the mod is deployed publicly to support further research in this field, offering a valuable testbed for future studies aimed at refining DDA techniques and enhancing player engagement.

Contents

1 Introduction

Action game is a popular type of video game all over the world. The key ingredient of this type of game that attracts players is fighting against opponents. While playing the game, players try their best to improve their skills, and upgrade characters, these all converge to a single objective: defeating the opponents.

Fighting in action games can come into two types of form: player versus player, and player versus non-player characters. Non-player characters(NPCs) are usually driven by scripted behavior. The purpose of these opponents is to provide a challenge to players, serving an enjoyable combat experience to keep players engaged. This approach allows game designers to ensure a fun and controlled experience for players within a limited timeframe. However, players' skills vary. A static scripted behavior makes it nearly impossible to deliver the same joy for players of different skill levels. While some players can only be engaged in hardcore gaming experiences like those found in Dark Souls[Fro16] and Elden Ring[Fro22], the other players prefer a more casual experience but reasonable challenge.

For commercial game developers, an ideal case to them is being able to access as much as gamers. Relying heavily on scripted behavior may not be sufficient to provide varied experiences for all players. One possible solution is offering multiple difficulty settings. However, it seems it is not enough. Take Helldivers 2[Stu24], a game released in 2024, as an example. Although the game provides 9 different difficulty settings for players to choose from, it finally leads to the same problem. Many players still struggling to find a suitable difficulty for them. This problem underscores the need for innovative approaches to difficulty that can adapt dynamically to the varied preferences and skill levels of modern gamers, which comes to dynamic difficulty adjustment(DDA).

Traditional dynamic difficulty adjustment methods in video games typically involve tuning various game parameters, such as enemy health, damage output, or the frequency of spawning enemies, to match the player's skill level. While these methods help maintain a balanced and engaging gameplay experience, they often fall short of capturing the complex behaviors shown by players.

In recent years, deep learning has presented an opportunity to enhance DDA through more effective techniques. By leveraging them, researchers can have a more in-depth understanding of players' skills, behaviors, and persona, creating adaptive systems that intelligently respond to different type of players. Meanwhile, considerable progress has demonstrated the feasibility of generating player-centered content or crafting agents capable of adapting themselves to players' skill levels. Research by Pfau et al.[PSM20] has investigated player experiences with deep-learning DDA agents, revealing a higher motivation and enjoyment compared to conventional DDA methods.

Build upon the achievements of past researchers. This research aims to address this challenge further by exploring the potential of DDA agents driven via deep player behavior modeling (DPBM) in an action rogue-like scenario, enhancing player motivation for repeated engagement by providing a dynamic and tailored challenge.

To reach this objective, I proposed the following research questions:

- How effectively does DPBM learn from previous runs within the context of a rogue-like game?
- Do players perceive the DPBM agent as accurately replicating their decision-making?
- Do players report high levels of player experience when facing the DPBM agent compared to the scripted one?

I hypothesize that the DPBM approach will successfully deliver a personalized gaming experience by capturing players' playstyles, thereby advancing research in DDA agent development. To evaluate this technique's effectiveness, I developed a mod, Dark Zagreus, for the award-winning action roguelike game, Hades[Gam20]. In this modified version, the final boss is replaced by a custom enemy opponent, which can learn from a player's most recent successful gameplay and replicate the decision-making of the player. The reason behind this adjustment stemmed from the final level's significance, where players have crafted their builds to confront the game's toughest challenge. While the original final boss has provided variability and excitement in testing diverse builds and strategies, its static behavior limits players' ability to fully explore and experiment with their unique strategies after multiple runs.

Over a two-week study period, 20 participants($n = 20$) took part in approximately one-hour sessions designed to verify the effectiveness of the DPBM approach in learning and replicating player playstyles by comparing the scripted agent and the DPBM agent. At the end of the study, a slight upward trend was discovered in all player experience metrics following the integration of DPBM. However, there was no significant difference being found, which means the upward trend can be interpreted as noise. For imitation effectiveness, a similar conclusion was found in the player experience. Aside from the quantitative result, some participants reported perceiving that their opponents used similar combat tactics to their own, which indicates the agent successfully captured player behavior in some cases. These results contribute to the field of dynamic difficulty adjustment and further explore the application of DPBM methods, offering the first step in applying DDA in action roguelike game scenarios.

After the study, I open-sourced the mod, which holds potential as an AI testbed for future research. This release aims to facilitate the development of more advanced DDA approaches. By making the mod available to the community, I expect to inspire further exploration and innovation in the field, ultimately enhancing DDA methods and improving player experience.

Figure 1. *The final boss fight in Hades.*

This master's thesis is structured as follows: Section 2 reviews prior research relevant to the topics addressed in this thesis, including player experience, dynamic difficulty adjustment, opponent imitation, and deep player behavior modeling. Section 3 introduces Hades, an action rogue-like game, which is the environment to perform the DPBM agent and conduct the experiment. Following this, Section 4 presents the mod I created, Dark Zagreus, and elaborates on the pipeline of the implementation of opponent imitation. Subsequently, Section 5 shows how the study for evaluation is conducted. The result of the survey will be shown in 6. In Section 7 and 8, the result is further evaluated and discussed. Finally, a conclusion is draw in Section 9.

2 Related Works

Difficulty in video games refers to the level of challenge presented to players. It consists of various aspects such as enemy strength, time constraints, and complexity of tasks. The right balance of difficulty is crucial for providing an engaging and rewarding experience for players. According to Flow Theory [Csi90], challenges that are too easy may become boring or lack excitement, while those that are too difficult may frustrate players and discourage further play. Achieving the optimal level of difficulty requires careful design and consideration of player skill levels, preferences, and progression throughout the game. However, it was almost impossible

to take care of players in various skill levels with a static challenge. Consequently, game developers and researchers continue to seek an efficient automated approach to difficulty tuning.

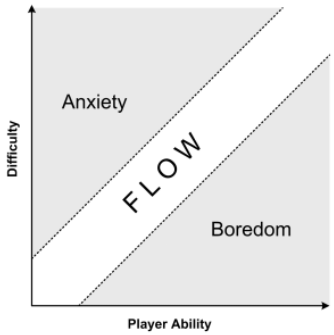


Figure 2. *A graph of flow theory[Csi90]. A well-balanced experience will be able to keep players in the flow, preventing anxiety and boredom.*

2.1 Dynamic Difficulty Adjustment

Addressing the challenge of covering players with varying skill levels, the gaming industry has increasingly focused on dynamic difficulty adjustment (DDA) as a key area of research. DDA aims to dynamically adapt a game's difficulty in real time, considering factors such as player performance, preferences, and proficiency. Extensive research[AM17][ASO13] has delved into the impact of DDA on player experience. These researchers found that players have higher engagement in DDA compared to a static setting, contributing significantly to this field.

For years, researchers have explored diverse directions to apply DDA, encompassing parameter tuning, dynamically generated environments, and game AI behavior modification. Parameter tuning, the earliest and most widely studied approach has seen significant development. For instance, Rhio et al.[Sut+15] implemented DDA in a tower defense scenario by dynamically adjusting enemy spawn rates and in-game rewards based on player performance metrics such as player remaining lives and enemy remaining health from previous levels. Pedro et al.[FJP21] explored generating game levels based on player personas to create personalized experiences. Notably, industry examples like Left 4 Dead 2[Cor09] demonstrate the practical application of DDA, where adjustments to enemy spawn rates and environmental elements are made in response to player performance.

In terms of game AI behavior modification, the role of AI behavior plays an important role in shaping the overall gaming experience. Even opponents with identical abilities can offer vastly different challenges to players based on their behavior and tactical decisions. For instance, a boss character having high damage output but predictable attack patterns may pose less of a challenge than one with dynamic

and unpredictable behavior.

To enhance player experience in DDA, a field of research has emerged that focuses on gaining a deeper understanding of player performance and playstyle.

2.2 Player Modeling

Effective DDA relies on the robust modeling of player behavior and abilities. Player modeling involves creating detailed representations of a player's skill level, preferences, and decision-making processes. This combines both implicit and explicit data collection methods. Implicit data includes in-game metrics such as reaction times, accuracy, completion rates, and preferred strategies. Explicit data can be gathered through direct player feedback or survey responses. Together, these data points allow developers to construct comprehensive profiles that can be used to tailor the gaming experience to individual players.

One popular method of player modeling is the use of player personas[Bar96] [CD09][Hol+21]. By categorizing players into distinct archetypes based on their behaviors and preferences, developers can create more personalized and engaging experiences. Dynamic difficulty adjustment systems can leverage these personas to adapt the game elements such as challenge level, narrative complexity, and interaction frequency, enhancing player satisfaction and retention. In recent years, deep learning approaches [dFF18] [Hol+14] have become popular to categorize player persona.

Opponent modeling focuses on understanding and predicting the behavior and strategy of an opponent. This involves learning the opponent's strategy and anticipating their next action.[NZ22] Researchers have been exploring the potential of opponent modeling in various genres of games.[SBS07][HB09][GS11][Wu+22] By utilizing opponent modeling, developers can analyze the opponent's skill level and find the best match to the player's skill level. Effective opponent modeling helps players better understand and adapt to the opponent's tactics and skill level in real time.

2.3 Opponent Imitation

In contrast to conventional AI research mainly focuses on devising optimal strategies for defeating players[Sil+16][Ber+19], the domain of DDA agents has a significant shift. By combining the advantage of player modeling and opponent modeling, researchers are able to create AI systems that mimic human-like behaviors and fit players' skill levels. This approach involves treating the player as the opponent and using advanced algorithms to predict and replicate player decision-making processes. Well-modeled AI opponents can make players feel more authentic and less

like traditional game bots. These agents integrate player modeling and opponent modeling techniques to customize AI behavior according to individual player persona or proficiency levels, creating an even game situation[PM21]. This approach has gathered considerable attention in diverse gaming genres, spanning from fighting games[Ish+18][Dem+17] to first-person shooters[KV21] and multiplayer online battle arenas (MOBAs)[SNC17]. Mirna et al.[SSC15] propose a method that prepares AI behaviors with different difficulties, changing the MOBA agent behavior based on the opponent's performance. By testing the agent against scripted agents and real players, their result shows that the approach can produce balanced matches in a high percentage of experiments. The evolving landscape of DDA agents showcases promising potential across various gaming environments, prompting ongoing exploration and research into their applicability in different game types.

Through the integration of machine learning techniques and player modeling, researchers have made significant strides in simulating player decision-making behavior. This progress has enlightened the development of agents aimed at emulating human-like gameplay. Previous studies have even demonstrated the creation of agents that are indistinguishable from real players by human observers[DP19][Pfa+20]. However, in the context of video games, the crucial factor is not only the believability and performance of these agents but also whether they contribute to an enjoyable gaming experience, whether as opponents or allies. The importance of player enjoyment has become increasingly apparent. For instance, Simon Demediuk et al.[Dem+19] investigated player enjoyment through encounters with various agents employing the Monte Carlo Tree Search technique. Surprisingly, their findings revealed that the most enjoyable agent to play against wasn't necessarily the one with the highest degree of realism.

To address this aspect and further enhance the player experience with DDA agents, Pfau et al.[PSM20][PSM18][PSM19] introduced the deep player behavior modeling (DPBM) method. Leveraging deep learning techniques, DPBM creates generative agents that successfully imitate player behavior. Unlike traditional approaches involving parameter tuning and scripted behavior, DPBM learns from players' decision-making patterns in past games, enabling DDA agents to emulate human opponents more effectively. In real-time, player and enemy states and player actions are compressed into input vectors and fed into a pre-trained neural network, which outputs probabilities for subsequent actions.

To evaluate player experience, Pfau et al. utilized an Intrinsic Motivation Inventory (IMI), assessing perceived competence, interest, tension, and effort. Their results convincingly demonstrated that DPBM agents offer a better player experience compared to the conventional heuristic parameter tuning approach. Ongoing advancements in this field continue to explore player experiences when engaging with believable agents[Moo+22][Lav+21][Roo+21], further refining the balance between realism and enjoyment in gaming encounters.

Given the absence of research on action rogue-like games and the exploration of fast adaptation models within a short playtime, this study aims to investigate the potential of employing a dynamic difficulty adjustment (DDA) agent to enhance player engagement in this genre. This master's thesis leverages deep player behavior modeling (DPBM), which has been proven effective in massively multiplayer online role-playing games (MMORPGs), and adapts it to the popular commercial game, Hades. The research focuses on evaluating player experience with the DPBM agent's emulation of their decision-making processes. The ultimate goal is to integrate dynamic AI behavior modification into action rogue-like games, thereby enhancing player immersion and enjoyment.

3 Hades

In this section, we introduce Hades[Gam20], the action rogue-like game used as the environment for developing and testing the DPBM agent. We provide an overview of the game's mechanics, its dynamic combat system, and the reasons for selecting Hades for this study.

3.1 Overview of Hades

Hades, developed by Supergiant Games, is a critically acclaimed action roguelike game that combines fast-paced combat with a compelling narrative. The game has sold over a million copies, indicating a large player group. In this game, the players will control Zagreus, the son of the Greek god Hades, attempting to escape from the underworld, fighting through many angry lost souls along the way. Unlike conventional games where progress drops when the character dies, Hades follows the roguelike tradition which uses death as a mechanic to push progress further.

A descriptive game loop is presented in Figure 3. To elaborate, the combats are happening within a run. A *run* refers to a single playthrough or attempt by the player to guide Zagreus, the protagonist, through the procedurally generated levels of the underworld. During a run, players navigate Zagreus through various chambers, each filled with enemies, traps, and challenges, with the ultimate goal of reaching the surface and escaping from the realm of Hades. A run begins when the player leaves the House of Hades and ends either when Zagreus successfully reaches the surface or when he is defeated in combat. Upon completion or failure of a run, players return to the House of Hades to regroup, upgrade Zagreus' abilities and weapons, interact with characters, and prepare for their next attempt.

This structure makes it particularly well-suited for this research project. A run typically lasts about 30 minutes, during which players can execute between 2000 to 3000 actions. This high frequency of actions within a relatively short time frame allows for collecting substantial data to extract a playstyle for training the DPBM

model. Additionally, the quick iteration between runs means players can rapidly test and observe the results of modifications, creating an ideal environment for iterative development and real-time feedback.

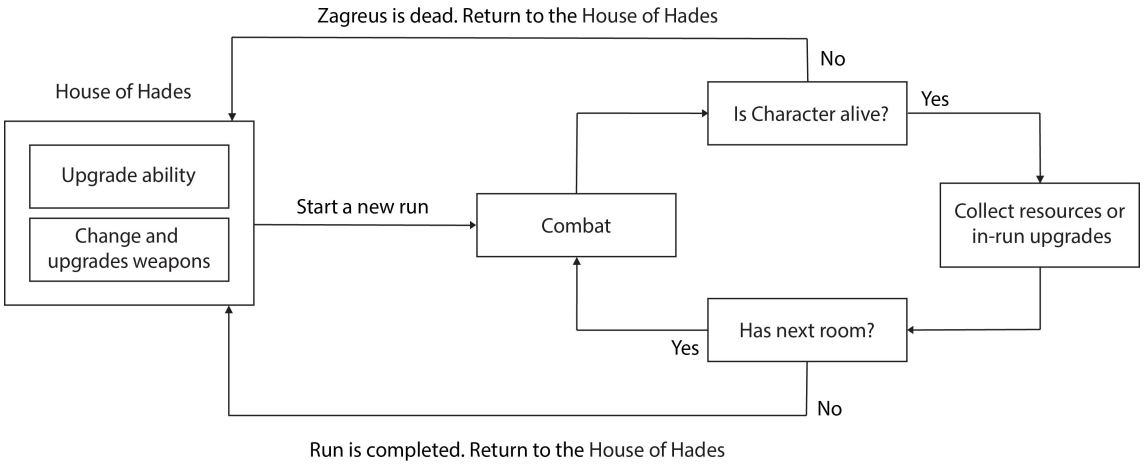


Figure 3. *The game loop of hades.*

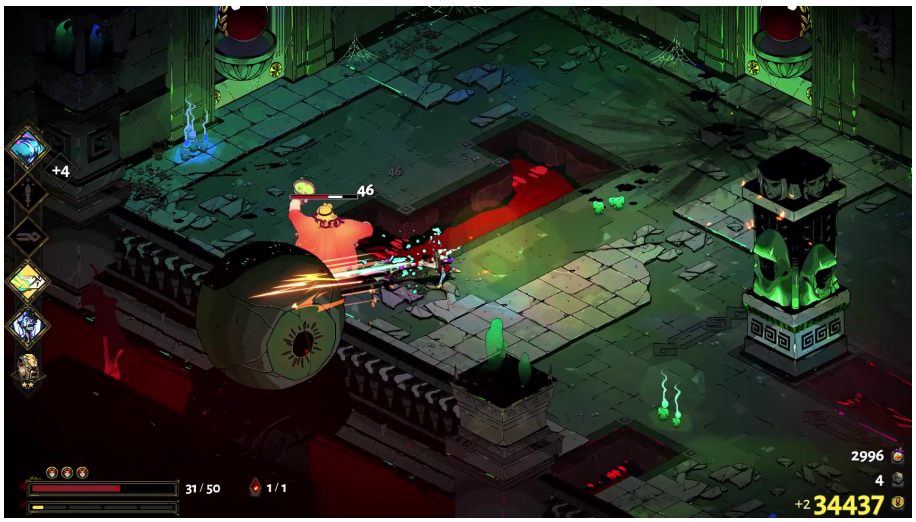


Figure 4. *Combat in Hades.*

3.2 Combat System

The combat in Hades is fast and dynamic, requiring players to utilize a combination of attacks, special abilities, dashes, and other actions to defeat enemies. The primary actions include:

- Movement: Players can move Zagreus around the game world using the directional controls (typically using a joystick or arrow keys).

- Attack: Players can perform regular attacks using a weapon such as a sword, spear, shield, or bow.
- Special Attack: In addition to regular attacks, players can unleash a special attack unique to each weapon. This attack often deals more damage or has special effects.
- Cast: Players can cast magical spells at enemies. These spells vary depending on the weapons equipped and boons acquired during gameplay. The default cast is shooting a red crystal that marks the enemy and increases the damage.
- Dash: Players can execute a quick dash movement to approach enemies, dodge enemy attacks, or navigate through traps.
- Call: Players can summon aid from powerful allies. These aids can inflict significant damage on enemies or provide defensive benefits.
- Summon: In addition to Calls, players can summon the assistance of various companions or entities throughout their journey. These summoned allies aid Zagreus in combat or offer other beneficial effects to aid in his escape from the underworld.

Figure 5. *A graph of available control in Hades.*

3.3 Weapons

Hades offers diverse types of weapons that players can use, each with a unique set of moves and playstyles. The primary weapons available in the game include the sword, bow, spear, fist, shield, and gun. Each weapon not only provides distinct

combat mechanics but also supports different strategies and tactics. For example, bow and gun weapons can reach enemies from a far distance while other weapons require players to be close to the enemy. Moreover, each weapon comes with four variants, known as Aspects, which further diversify their capabilities. These variants offer unique moves, abilities, and enhancements, allowing players to tailor their combat experience to their preferences and playstyle. This variety in weapons and their Aspects contributes significantly to the depth and replayability of Hades.

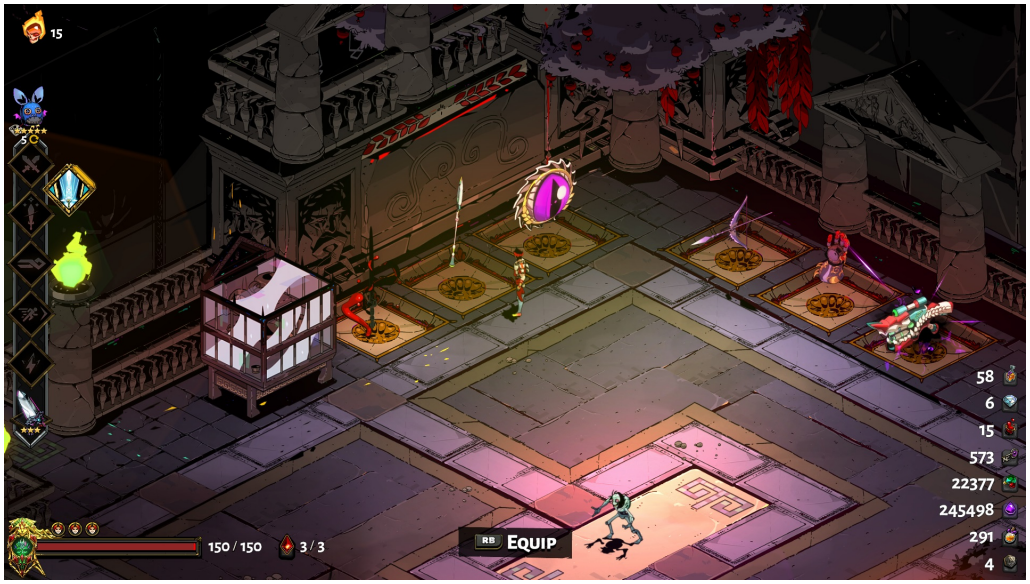


Figure 6. *A picture of Arsenal, where players can choose from a variety of weapons available in the game, each offering a unique playstyle and set of abilities.*

3.4 Upgrades

In addition to weapons, players can further enhance Zagreus through various upgrades. Permanent upgrades can be obtained in the House of Hades, providing lasting benefits to Zagreus. The Mirror of Night, located in Zagreus's room, offers permanent abilities that reduce the chance of dying during a run. Additionally, players can upgrade their weapons' levels, increasing damage or enhancing weapon abilities.

Within a run, players can acquire temporary upgrades by winning combat encounters in different chambers. These upgrades, which reset after each run, come in two forms: Boons and Daedalus Hammers. Boons are blessings from various Olympian gods such as Zeus, Athena, Artemis, and Poseidon. These boons grant Zagreus powerful abilities, buffs, and enhancements that can significantly alter gameplay. For example, Zeus might grant the ability to summon lightning bolts on enemies, while Athena could provide a shield that deflects attacks. Players select boons

Figure 7. *An example of weapon variants in Hades.*

Figure 8. *A picture of the Mirro of Night, which is the interface for players to improve Zagreus's ability. The upgrades will permanently exist.*

as they progress, allowing them to tailor their abilities to their preferred playstyle.

The Daedalus Hammer is another upgrade that players encounter during runs. It enhances weapons by modifying attack patterns, adding new abilities, or increasing damage output. For instance, a Daedalus Hammer upgrade might increase the attack speed of Zagreus' sword or increase the number of arrows shot from special attacks of bows.

These mechanics ensure that each run is diverse and unique, offering different builds and strategies for players to explore.

3.5 Enemies

In the combat of Hades, players confront two primary categories of adversaries: the minions, representing the common threats, and the bosses, comprising strong adversaries encountered at pivotal junctures. The minions encompass a variety of foes, ranging from the basic lost souls that move very slowly to the agile souls that hold dangerous weapons. These adversaries present diverse challenges, utilizing melee strikes, ranged attacks, and unique abilities to impede Zagreus' progress. Meanwhile, bosses, such as the Furies and other powerful adversaries, serve as significant obstacles, demanding advanced tactics and mastery of combat mechanics to overcome. Throughout the journey, players must navigate encounters with both minions and bosses, employing strategic prowess and skillful execution to escape the underworld.

Figure 9. *Minions in Hades.*

3.6 Modding Environment

Hades offers a robust and accessible modding environment, making it an ideal platform for integrating a DDA agent. Most of the gameplay code, including combat mechanics, upgrades, levels, and weapons, is available in Lua. This eliminates the need for extensive reverse engineering, making the modding process significantly more straightforward and user-friendly. Additionally, Hades has a vibrant modding community that provides extensive tutorials and tools, facilitating the modding process and offering resources and guidance to help modders navigate and manipulate

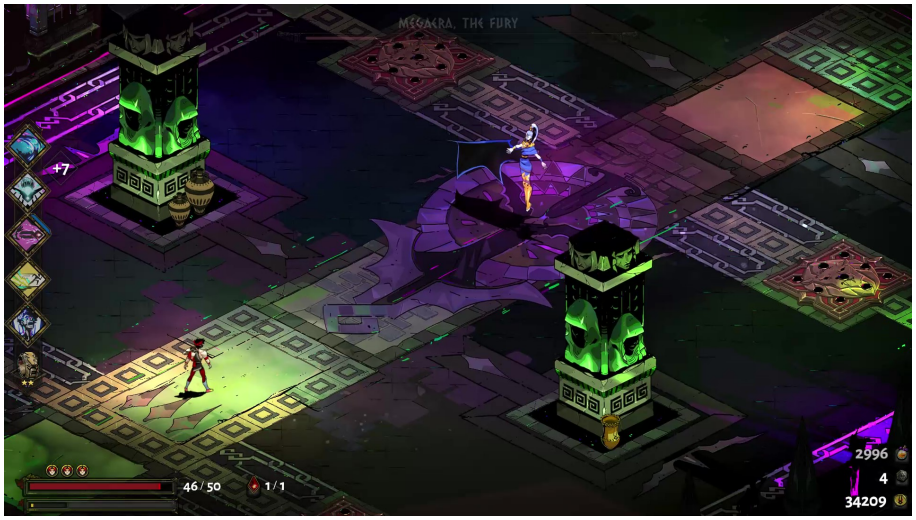


Figure 10. *A picture of a boss fight with Megaera, the Fury.*

the game's code effectively.

One of the key advantages of Hades' modding environment is the accessibility of the enemy AI code, allowing me to directly modify it with my own AI scripts. This capability is crucial for integrating the deep learning model, enabling me to tailor enemy behaviors to mirror player strategies and tactics.

Despite its modding-friendly advantage, there are some restrictions of its modding environment that create challenges. Firstly, it is an enclosed environment that does not allow importing external libraries. This limitation posed a challenge for implementing a deep learning model, however, I managed to overcome it, which will be explained in Section 4. Secondly, certain Lua native modules, such as io and os, are not available due to security concerns. Allowing these modules could enable modders to access and manipulate players' devices inappropriately. For example, the io module could be used to read, edit, or write files on players' devices, potentially leading to security breaches. This becomes an obstacle for me to save players' gameplay data properly. Luckily, Hades allows modders to save data alongside the game state. This feature is particularly useful for saving data for training purposes, making the goal more accessible.

Hades can be launched in three different versions: DirectX, Vulkan, and 32-bit. The DirectX and Vulkan versions are 64-bit and are the default, more commonly used versions. The 32-bit version, however, is more conducive to debugging because it allows the use of io and os modules. This flexibility in the 32-bit version aids in development and testing. Furthermore, the modding community has developed tools to extend modding capabilities, such as enabling the use of Python. However, requiring players to install Python to use a mod can introduce installation issues and decrease player interest. To maintain accessibility, I avoid using the io and

os modules and the additional Python tools. However, this feature still becomes a useful tool to export gameplay data and evaluate the DPBM model externally.

To conclude, although there are some restrictions, they do not affect the research significantly. Given these benefits, Hades is an excellent fit for integrating a DDA agent. The accessible gameplay code, supportive modding community, and ability to modify enemy AI provide an ideal platform for this project.

3.7 Project Scope

While Hades offers a wide range of actions, I decided to exclude Cast, Call, and Summon from this study due to their system complexity, as these features are specifically designed for players and are challenging to replicate. This decision was made after careful consideration of the system's intricacies and the time constraints of this research. Similarly, Boons and Hammer upgrades were also excluded due to their complexity. Despite missing these significant combat strategy elements, I believe that the DPBM approach can still capture the majority of player behavior since the most commonly used actions were implemented.

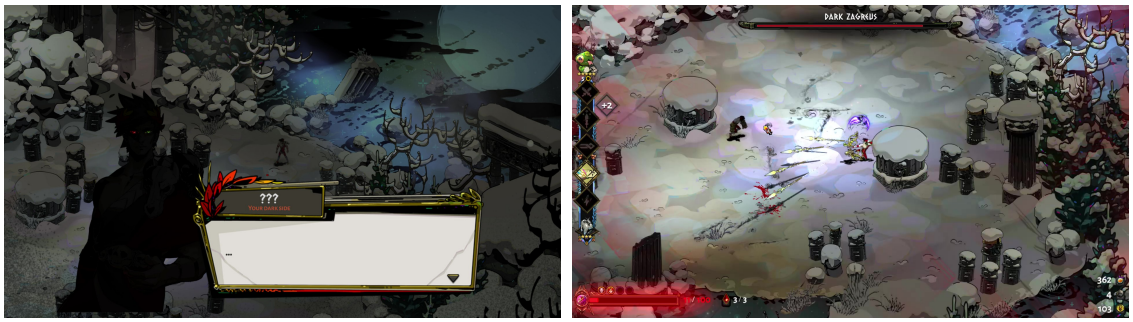
4 Implementation

In this section, the mod, Dark Zagreus, is introduced as the integration for the DPBM agent. Next, the pipeline of the implementation is provided. The approach involves gathering player data from the most recent successful run and training a deep-learning model in real time when the player enters the boss room. This ensures that the agent is adapted to the player's most recent strategy and behavior. Each model is independently trained from the data collected in the last successful run, accommodating the varying strategies players may employ based on the resources they have within each run.

4.1 Dark Zagreus

Dark Zagreus is the mod created for this research, replacing the final boss, Hades, with a mirrored version of the player character. Players will encounter Dark Zagreus when they have a complete build for their combat strategy, making it a critical point to test their skills. Dark Zagreus serves as the enemy integrated with the DPBM adaptation. During the first encounter, Dark Zagreus is driven by static behavior, providing a baseline challenge for the player. Upon the player's first successful defeat of Dark Zagreus, it will learn from the run, adapting its tactics based on the player's victorious strategy. I specifically exclude runs where the player dies, as the strategies used to successfully defeat the boss are more valuable and interpretable for the learning model.

To enhance the immersive experience, the dialogues of the final boss and the name within the game content are replaced to fit with the content, shown in Figure 11.



(a) The dialogues in the final boss fight are replaced. *(b) The player fighting against Dark Zagreus.*

Figure 11. *The mod, Dark Zagreus.*

4.2 Pipeline Overview

The implementation pipeline of the mod is divided into three main stages: data collection, deep player behavior modeling, and DDA agent integration. Each stage is managed by a specific system: the logging system for data collection, the player modeling system for behavior modeling, and the enemy AI system for DDA agent integration. The following subsections provide a detailed description of each stage. All of the systems are written in Lua, which is the scripting language used in Hades modding.

The pipeline is designed to learn from the player's gameplay and adapt the AI opponent accordingly. The process is illustrated in Figure 12.

When a new run starts, the logging system is activated. As the player progresses through the run, the system is triggered by specific actions of interest. For each such action, the system logs a row of data containing the current state of the character and the decision made by the player. If the player dies during the run, the recording is terminated, and no data is retained.

While entering the boss room, the player modeling system checks if data from the previous successful run exists. If such data is available, the system trains a new model using this data. If no previous successful run data is available, the AI agent is set to a default mode.

The boss fight then runs with the DDA agent controlled by the trained model (if available) or operating in default mode. Once the boss fight ends, the outcome determines whether to save the recorded data. If the player survives the boss fight, the recorded data is saved for use in the next run. If the player dies, which indicates

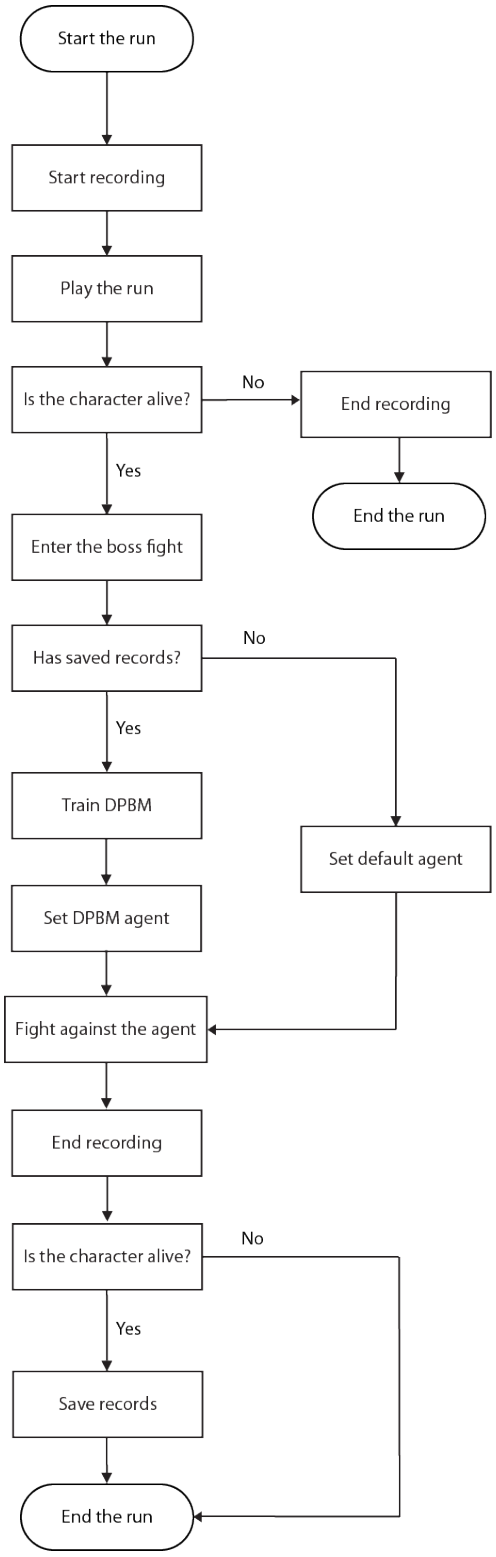


Figure 12. *The pipeline overview.*

a failed strategy to complete a run, the data is erased. This pipeline ensures that the AI opponent is updated to reflect the player's most recent tactics and strategies, providing a tailored and dynamic challenge in each run.

4.3 Data Collection

As stated in the overview, data collection is the first stage in the pipeline. This stage involves implementing a data logging system within Hades to capture players' gameplay data during each run. The primary goal is to gather detailed information on the player's decision-making based on the game state, creating a dataset for training the model.

To achieve this, I designed a logging mechanism that activates at the start of each run. The system continuously observes the player's gameplay and is triggered when the player performs specific actions of interest. Specifically, the system listens for weapon fire events. When a weapon is fired, the system logs the action into a queue, subsequently pushing this record from the queue into storage. The system overview is presented in Figure 13. The actions of interest include attack, special attack, dash, and reload, which are the most commonly used actions in combat (as shown in Table 1). Note that, in Hades, dashing is also considered a weapon. While other actions like Cast, Call, and Summon also affect player combat strategy, I decided not to take them into account due to the system complexity of these three actions and the time constraint of this research.

The use of a queue is essential because some weapons feature an advanced attack that requires holding the attack button. To accurately log these actions, the attack action is distinguished between regular and advanced attacks. When the player performs an advanced attack, the character first executes a regular attack before starting to charge the advanced attack. To avoid logging the regular attack triggered by the advanced attack, the system logs the regular attack into the queue and then overrides it with the advanced attack. This ensures that only the advanced attack is recorded, providing an accurate representation of the player's decision-making.

Additionally, through several playtests, I identified that dashes can be categorized into two distinct types: dashing toward and dashing away from enemies. These categories reflect different combat strategies. For example, players using melee weapons often dash toward enemies to close the distance, whereas players with ranged weapons use dashes to maintain distance from their targets. Which type of a certain dash is distinguished based on whether the difference between the angle toward players and the angle toward the fire direction is larger than 90 degrees.

Notably, I found an outlier occurring when the player defeats an encounter. Before entering the next room, there is a free session for the player to perform

any actions and upgrade their abilities. To address this, I excluded these actions performed by verifying if the room contained an enemy. After implementing this validation, a significant increase was observed in the model's accuracy.

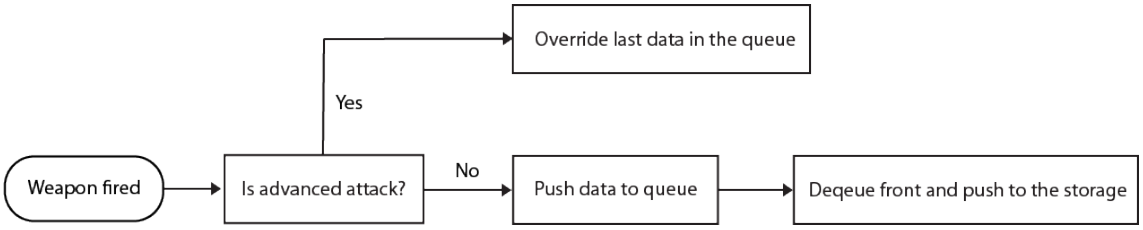


Figure 13. *The workflow of the recording system.*

Actions	
Attack	Regular attack of a weapon
AdvancedAttack	Advanced regular attack of a weapon
SpecialAttack	Special attack of a weapon
DashToward	Dash toward closest enemy
DashAway	Dash away from closest enemy
Reload	Reload a gun weapon

Table 1. *Types of action that trigger the recording system.*

Alongside these actions, the logging system also captures the game state each time an action is recorded. This is crucial because understanding the context in which a player makes a decision provides deeper insights into their behavior. The logged states include detailed information about both the player's state and the state of the closest enemy, as listed in Table 2. The decision of choosing to focus on the closest enemy is because, unlike in boss fights, players often encounter multiple enemies within a level, making it difficult to identify which enemy the player is targeting. By focusing on the closest enemy, the player modeling system can reasonably infer the player's primary opponent.

Before a record is logged, all the elements are first normalized for consistency. Normalizing the data ensures that all features are on a similar scale, which is crucial for the efficient training of the model. Specifically, the health values are normalized by dividing by the maximum health of the character, while distances are normalized by dividing by 1000, which is the maximum range that the player can reach with the farthest-ranged weapon.

Additionally, the system captures whether the player deals damage to an enemy or receives damage, as these events can significantly influence decision-making. For instance, a player might immediately dash when taking damage or become more

States	
OwnHealth	Player's health
EnemyHealth	Closest enemy's health
Distance	Closest enemy's distance
GetDamagedRecently	Is damaged within 1 second
DamageEnemyRecently	Damage an enemy within 1 second
HasActiveBuff	Is special power activated
IsReloading	Is player reloading the weapon
Ammo	The amount of ammo loaded

Table 2. *Types of state that captured by the recording system.*

aggressive after successfully damaging an enemy. These factors are logged as binary values, with 1 indicating true and 0 indicating false.

To capture players' strategy of using specific weapons, whether any buffs are activated is added to the log records. Some weapons have special abilities that enhance the player. Take the Aspect of Chiron (a variant of the bow weapon) as an example, when a regular attack hits the enemy, it will mark the target. With this enhancement activated, the special attack can fire multiple arrows that track the target. To log this, similar to the damage factor, it is logged as 1 for activated or 0 for deactivated.

Finally, information specific to gun weapons is captured, such as whether the player is reloading and the amount of ammo available. Reloading temporarily disables weapon firing, forcing the player to rely on dashing, while the amount of ammo can affect the player's tactical decisions, such as whether to reload when ammo is low. The reloading status is recorded as 1 or 0, and the ammo count is normalized by dividing by the maximum ammo capacity of the weapon.

By combining actions with their corresponding states in each logged entry, the player modeling system can thoroughly analyze the conditions under which players make specific decisions, thereby gaining a deeper understanding of their behavior and strategies.

To handle the data generated by the logging system, the data is initially stored in memory. Due to environment restrictions mentioned in Section 3, custom scripts are not allowed to use the io module in the regular version of the game. Consequently, the data is then saved with the game state when the run ends, which is then automatically saved with the save file. Since the player modeling system is only interested in the data from the most recent successful run, the old data is overwritten by the latest run, preserving the saved file size.

The data schema is designed to efficiently organize and manage the collected

gameplay data. The schema begins with metadata, followed by detailed logs of player actions and corresponding game states. Specifically:

- Version: The first row indicates the version of the data schema, which increments whenever there are changes to the schema.
- Weapon: The second row records the type of weapon the player uses.
- Weapon Index: The third row indicates the index of the specific variant of the weapon, with each type having four variants.
- Action Logs: Starting from the fourth row, the logs from the logging system are recorded. The even rows contain the state of the game when an action is taken, and the odd rows contain the action itself, representing the decision made based on the preceding state.

This structured approach allows maintaining a clear and organized dataset, facilitating subsequent data analysis and model training.

To accommodate testing and debugging, I leveraged the fact that Supergiant Games allows the io module in the x86 version of Hades, which is selectable on some platforms. This enables the opportunity to develop a feature that exports the recorded data to external files for testing purposes. By utilizing the x86 version for development, the data collection system can be debugged and refined more effectively, and the collected data can be further examined externally.

4.4 Deep Player Behavior Modeling

Building on the data collected from player actions and game states, the next stage in the pipeline involves modeling player behavior using deep learning techniques. By training on the dataset generated in the data collection stage, the model learns and predicts player actions based on the game state.

To accurately model player behavior, I designed a neural network architecture that balances complexity and performance. The architecture consists of four layers. The input layer, with 20 neurons, takes the game state and the past actions performed by the player as input. Next, this is followed by two hidden layers, each with 13 neurons, composed of dense layers with sigmoid activation functions to identify patterns and relationships in the data. The output layer, with 6 neurons, generates the predicted action as a probability distribution over the six actions mentioned in Section 4.3: attack, advanced attack, special attack, dash away, dash toward, and reload. The model architecture is illustrated in Figure 14.

Before training the model, the collected data is preprocessed to ensure suitability for input into the neural network. Inspired by the study by Pfau et al. [PSM20], I included the last two actions performed by the player alongside the game state to

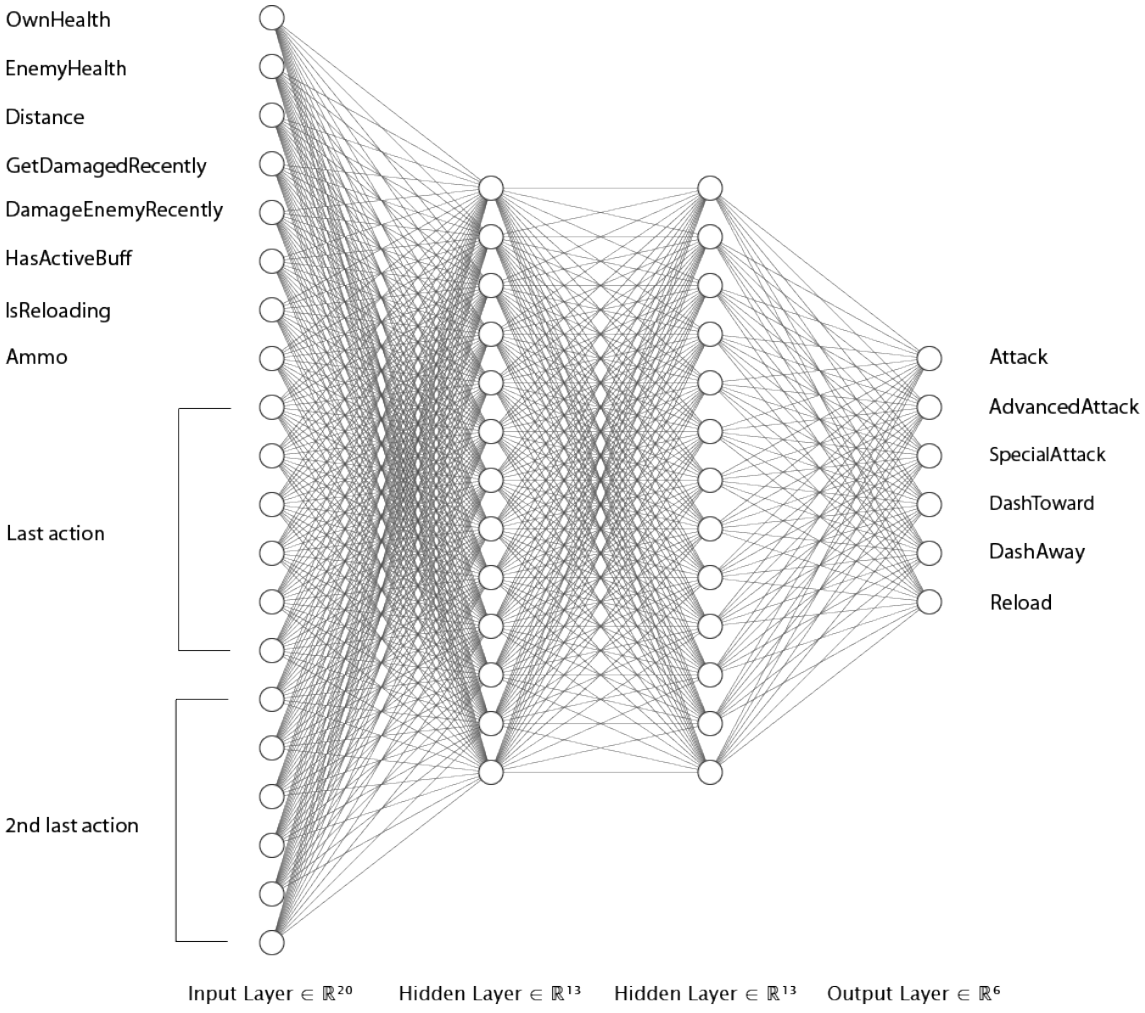


Figure 14. *The model architecture of DPBM.*

capture potential patterns or weapon combos that players tend to use. The reason only two actions are captured is because of the notable increase in training time with capturing three past actions. Since the first and second rows of records do not have the last two actions, these rows are excluded from the dataset after preprocessing.

The model is trained using supervised learning, where input game states are mapped to the corresponding player actions. Due to the environment constraints of Hades, which limit the ability to import external machine learning libraries, I used an open-source lightweight neural network implementation[wix13] in Lua by Wixico on GitHub, with sigmoid as activation function and mean square errors(MSE) as the loss function. This implementation was then modified to fit the modding environment of Hades. The training process involves running backpropagation through the dataset for a fixed 10 epochs, adjusting the model weights to improve prediction accuracy. The learning rate is set to 0.03. Since Hades does not support background threading, the training process occurs on the main thread, necessitating a short training time to avoid disrupting the game loop and affecting player experience.

To assess the performance of the model, a debug feature is developed for exporting the recorded data and evaluating the model externally. A custom evaluation metric is introduced to assess prediction accuracy. For each prediction, the action with the highest probability is compared to the actual player action. The accuracy is then calculated by averaging the correct predictions over the total predictions, resulting in a percentage accuracy score. It is calculated from 5-fold cross-validation externally. The validation process splits the dataset into five parts and uses one part as the validation set in each of the five iterations. The validation set, which is not used during training, provides an unbiased evaluation of the model's performance. Next, the accuracy scores of the five iterations are averaged as the final result. By analyzing this metric, I fine-tune the model parameters and architecture to achieve optimal results and ensure a reasonable training time.

4.5 DDA Agent Intergration

The final stage in the pipeline is the integration of the trained DPBM into the game environment of Hades.

Integrating DPBM into Hades involves several critical steps to ensure seamless functionality and an engaging player experience. First, I ensure that the DDA agent can access and utilize the trained model during gameplay. This is achieved by implementing an opponent character, Dark Zagreus, who mirrors the main character's appearance but with darker color animations for easy identification by the player.

To maintain the game's balance and avoid crashes, a set of custom weapons are implemented. These weapons replicate the player's weapons and their variants for Dark Zagreus to use, ensuring they function identically to the originals. The rea-

son for replicating the weapons set on the original weapons are specially designed for player characters, and using them directly for enemies would cause the game to crash. In addition, by using custom weapons, the damages of the weapons can be decreased for a better adjustment for balancing since the original weapons are designed to defeat enemies, who have relatively high health points.

When the player enters the boss room, a validation is performed to check if there is a previous gameplay record saved alongside the game state. If there is no record, the DDA agent equips the initial weapon of the main character and uses static probabilities for actions. If a record exists, the agent switches to DPBM mode and loads the trained model.

Once the model is loaded, the model operates in real-time, making decisions based on the current game state. To seamlessly integrate DPBM's decisions with Hades' existing game mechanics, I developed an enemy AI script that translates model outputs into in-game actions. This script ensures that Dark Zagreus can perform actions such as attacking, dashing, and using special abilities in a manner consistent with the game mechanics.

In the initial stages, Dark Zagreus performs random actions due to the absence of past data. As the game progresses and the agent accumulates data from previous actions, the model begins to make decisions. The decision-making process considers the current state and past actions, producing corresponding probabilities to determine the next action. The agent then randomly selects an action based on these probabilities. This method ensures a variety of actions and prevents the agent from repeatedly performing the same action, even if one action has a slightly higher probability. This approach effectively replicates the player's playstyle while maintaining a dynamic and unpredictable AI opponent.

5 Study

To evaluate the effectiveness of the approach, a study long lasted two weeks was conducted. Participants were asked to play Hades with the mod integrated into the game and report their experiences of fighting against the DDA agent before and after it adapted to their behavior within an approximated one-hour session. To reach a broad audience, the mod was published on a popular website for game modifications and distributed the publication message to several communities.

Within the study, participants were required to play two runs of Hades with the mod and fill out a questionnaire, with each run lasting roughly 30 minutes and the questionnaire taking about 10 minutes, totaling 1 hour and 10 minutes. The study was conducted over 2 weeks. Participants first provided consent and were informed that no personally identifiable information would be collected.

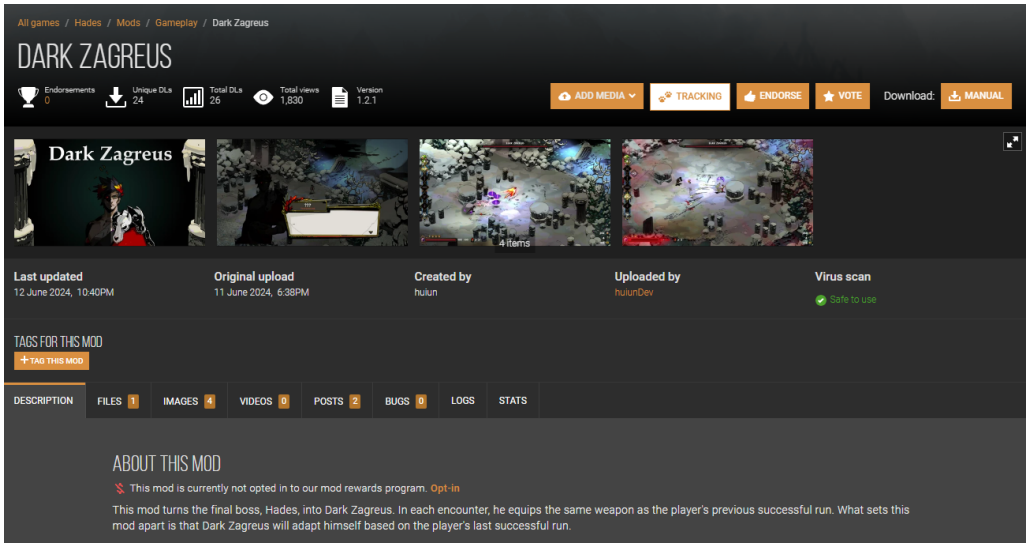


Figure 15. *The mod page of Dark Zagreus.*

In the first run, participants played Hades with the goal of defeating the final boss. If they lost, they were required to play again until they reached the final boss. The first final boss they encountered was controlled by a default AI agent with static probabilities for actions. After this playthrough, participants answered a set of questions about their experience with the AI agent.

In the second run, the goal was to reach the final boss level, with defeating the boss being optional. This time, the AI agent used data collected from the first playthrough to train the DPBM agent. Upon entering the final boss room, the AI adapted to the player's behavior from the first run. If participants did not reach the final boss, they had to restart the second playthrough. After completing this run, they answered a second set of questions about their experience. Both sets of questions included options to upload anonymous play records and leave comments about their experiences with the AI agent. The questionnaire concluded with an optional section for suggestions about the mod.

5.1 Procedure

To reach a broad audience, I published the mod on Nexus Mods[Mod]. The deployment message was majorly distributed to the Hades modding community[Com] and Hades subreddit[Red], these two sites contained a large amount of Hades players viewing every day. Within the publication message, a detailed installation tutorial and detailed instructions for participation were provided to players.

The detailed steps for the study were as follows:

- Consent and Information: Participants provided consent and were informed

about the study's details and the confidentiality of their data.

- First Run: Participants played the first run of Hades with the goal of defeating the final boss. The initial AI was controlled by a default behavior model.
- First Questionnaire: After the first run, participants answered questions about their experience with the initial AI agent.
- Second Run: Participants played the second run, aiming to reach the final boss. This time, the AI adapted to their playstyle using data from the first run.
- Second Questionnaire: After the second run, participants answered questions about their experience with the adapted AI agent.

During the study, I concealed the fact that Dark Zagreus was driven by a deep learning model to avoid bias, instead providing a vague description: "After you defeat Dark Zagreus, he will return to the shadow and observe how you play the game and adapt himself to your play style."

5.2 Participants

The criteria for participants were straightforward: they had to be players of Hades, regardless of their skill level, and have the game available on Windows or macOS. At the end of the study, the mod received over 1700 views, 100 people took part in the study, 33 of them finished the first playthrough and 20 participants (n = 20) ultimately completed the questionnaire.

5.3 Measures

The questionnaire was conducted using Qualtrics, the official survey tool of Utrecht University. To measure player experience, items from the mini Player Experience Inventory (miniPXI)[Hai+22] were used, focusing on challenge, autonomy, curiosity, immersion, mastery, meaning, and enjoyment. Participants responded to questions on a 5-point scale, ranging from "strongly disagree" to "strongly agree," which were rescaled to a range of 1 to 5 for analysis. Additionally, a qualitative question, which was also scored from 1 to 5, was included for asking whether fighting the final boss felt like combating themselves to assess if the DPBM approach successfully imitated the player's behavior. These scores were then compared and calculated p values with a two-tailed repeated measures t-test. At the end of each set of questions about the two runs, participants were also asked to upload their play records, which is optional. These play records were used to quantitatively assess the accuracy of the model.

6 Result

After the study period ended, results were calculated for the miniPXI items. As shown in Table 3, a majority of median scores for the items saw an increase after applying the DPBM adaption. However, no significant differences ($p < 0.05$) were found between the before and after conditions for any of the items while using a two-tailed repeated measures t-test.

Metric	Mean Before	Mean After	Median Before	Median After	Std Before	Std After	p_value
Challenge	3.00	3.20	3.00	3.50	1.30	1.47	0.64
Autonomy	4.20	4.60	4.50	5.00	1.06	0.68	0.09
Curiosity	4.45	4.65	5.00	5.00	1.00	0.59	0.26
Immersion	4.35	4.20	5.00	5.00	1.04	1.11	0.65
Mastery	3.00	3.40	3.00	3.50	1.45	1.43	0.20
Meaning	3.50	3.75	3.50	4.00	1.24	1.16	0.33
Enjoyment	3.75	4.05	4.00	4.00	1.21	1.15	0.28

Table 3. *Summary of miniPXI metrics Before and After AI Adaptation*

To look more in detail, among the player experience metrics, several key items showed slight trends. Challenge and Mastery are crucial as they directly relate to the player's sense of progress and skill development. The median score for Challenge increased from 3.00 to 3.50. This increase aligns with a sense of Mastery, which also saw its median score rise from 3.00 to 3.50.

As for Autonomy, which reflects the players' perception of control and freedom within the game. The median score increased from 4.50 to 5.00.

Notably, Curiosity and Immersion stay at a high score in both conditions. Enjoyment is a comprehensive measure of the overall gaming experience. Although the median score for Enjoyment remained constant at 4.00.

Aside from the miniPXI metric, the Imitation score, which assesses whether players felt Dark Zagreus mirrored their playstyle, showed increases from a mean of 2.90 to 3.65 and a median of 3.00 to 4.00.

Mean Before	Mean After	Median Before	Median After	Std Before	Std After	p_value
2.90	3.65	3.00	4.00	1.41	1.57	0.16

Table 4. *Summary of Imitation score*

In addition to the quantitative measures, several open-ended responses were collected from participants regarding their experiences fighting against Dark Zagreus before and after the DPBM adaptation. Out of the 20 participants, 18 provided feedback on their encounters with the default AI agent. Of these responses, 33.3% were positive about the enjoyment of the fight, with comments such as "it was fun to fight against.", while the others stayed neutral. 5 of the 18 responses indicated

that the default AI was challenging, whereas 8 responses described it as too easy to defeat, with remarks like "the fight itself was kind of easy, just took a bit to learn the attack pattern" and "awkward". The rest of the responses offered suggestions without specifically mentioning enjoyment or difficulty.

After the agent adapted its behavior using the DPBM, this study received 16 open-ended responses about the participants' experiences. Notably, 37.5% of the responses mentioned they observed that the agent mimicked their behavior, with quotes such as "It felt quite sad to see my poor skills reflected in Dark Zagreus," "I saw him taking advantage of special attacks like me. It's kind of interesting," and "Definitely felt like I was fighting against my own scummy tactics." Conversely, 18.8% of participants mentioned that they did not feel the agent replicated their play style, indicating some variability in the perceived effectiveness of the DPBM. Moreover, 2 participants highlighted the increased challenge and value of the adapted AI, with one notable comment stating, "a hard but valuable challenge that helped me improve the run after.", while 7 responses described the fight as very easy, with one participant commenting, "I feel like it's too easy to defeat him. If I want, he won't even have any chance to hurt me." The rest of the responses stay neutral.

For the model accuracy, 11 gameplay records were uploaded by the participants voluntarily. The analysis result is shown in Figure 16 and Table 5. The model achieved a mean accuracy of 55% when trained using an average of 2672 data points. When examining the accuracy for modeling player behavior with each weapon type, guns demonstrated the highest accuracy with 68%, while swords showed the lowest with 44%. Swords had the highest action count of 3147 and the standard deviation showed a high variety. Spears, with less training data of around 2350, had relatively lower accuracy compared to bows and guns, which had similar training data counts of around 2600.

Weapon	Sample Count	Mean Action Count	Std Action Count	Mean Accuracy	Std Accuracy
Bow	2	2628.5	111.016	0.59	0.014
Gun	2	2658.5	560.736	0.68	0.070
Spear	4	2343.75	718.290	0.56	0.112
Sword	3	3147.667	1052.905	0.44	0.046
Fist	0	0	0	0	0
Shield	0	0	0	0	0

Table 5. *Summary of DPBM for each weapon*

With the imitation score and DPBM accuracy calculated, the correlation between these two metrics was further investigated. Although there were 11 gameplay data sets, only 5 were recorded during the first playthrough and could be compared with the imitation score in the second playthrough. Pearson's correlation coefficient was used to measure the correlation. The result showed a correlation coefficient of -0.80 and a p-value of 0.10.

Figure 16. *The prediction accuracy of DPBM.*

Weapon	Imitation Score	DPBM Accuracy
Gun	4	0.73
Sword	5	0.49
Gun	4	0.63
Sword	5	0.43
Bow	5	0.60

Table 6. *Imitation perception score and DPBM prediction score of participants that upload the first-run record.*

7 Discussion

The results show valuable insights into the effectiveness and impact of integrating dynamic difficulty adjustment with DPBM into Hades. While the quantitative analysis did not reveal statistically significant differences in the player experience metrics before and after the DPBM adaptation, several slight trends still emerged.

Firstly, in terms of the miniPXI metrics, the increase in median scores for Challenge and Mastery suggested that the DPBM agent introduced a more challenging and skill-testing environment. Players perceived the game as more demanding, which aligns with the goal of creating a dynamic and engaging experience. This increased difficulty likely contributed to a greater sense of accomplishment and skill development, enhancing overall player engagement.

Curiosity and Immersion remained consistently high in both conditions, which showed that the difficulty adjustment did not downgrade the game's immersive qualities. Players remained interested and invested in the game, eager to see how the DPBM agent would respond to their actions.

Correlation Coefficient	-0.80
p_value	0.10

Table 7. *Correlation between imitation perception score and DPBM prediction accuracy.*

Similarly, the unchanged enjoyment scores reflected a high level of player satisfaction. This suggests that this approach maintained the game's entertainment value, ensuring that the modifications did not negatively impact the overall gaming experience.

Like the miniPXI metrics, a slight increase was observed in the Imitation score, where players felt that the DPBM agent more closely mirrored their play style. This perception of opponent imitation as a reflection of their strategies contributed to a more personalized and engaging experience. The qualitative feedback further supported this finding, with several participants explicitly noting that DPBM mimics their tactics. However, it comes to the same conclusion as the player experience, which is not enough to make a bold statement.

In contrast to the supportive result regarding imitation score, there is variability in the perceived effectiveness of the DPBM, with some participants not feeling their playstyle was accurately replicated, suggesting areas for further refinement. The feedback pointing to the increased challenge and value of the DDA agent highlights the potential for the DPBM to enhance not only the immediate gaming experience but also players' skill development over time.

The DPBM's accuracy, averaging 55% across 11 gameplay records, is significantly higher than the expected 17% accuracy if action probabilities were evenly distributed among the six possible actions. This performance is comparable to previous work[PSM19][PSM20] on DPBM in other scenarios, which achieved a mean accuracy of 70.3% and 60.64%. A further comparison is presented in Table 8. Considering that the current DPBM model accounts for fewer factors in a complex environment, this substantial accuracy suggests that the DPBM effectively learns and applies player behaviors.

	Game	**Genre**	**DPBM Input**	**DPBM Output**	**Mean Accuracy**	**Std Accuracy**
This work	Hades	2D Action Game	20	6	55.0%	11.0%
Pfau et al.[PSM19]	Custom Testbed	2D Fighting Game	24	9	70.3%	13.5%
Pfau et al.[PSM20]	Aion	3D MMO RPG	Avg 98.2	Avg 76.2	60.64%	22.57%

Table 8. *Comparison of my DPBM approach with related work*

Upon examining the accuracy of each weapon type, several interesting insights emerged. Despite having relatively few data points, the prediction accuracy for guns was high, suggesting that the strategy for using guns might be relatively straightforward. In contrast, swords, which had the most data points, showed lower accuracy,

indicating that the strategy for using swords might be more complex. However, due to the small sample size, these findings require further investigation with a larger dataset.

Regarding the correlation between imitation scores and DPBM accuracy, the data sample was too limited to yield significant results. Therefore, further research with a larger sample size is needed to draw more definitive conclusions.

Overall, these findings suggest that while the DPBM adaptation did not result in statistically significant changes across all metrics, it introduced meaningful enhancements to player experience, which responds to the research questions:

- The quantitative analysis of DPBM accuracy indicates a high efficiency of learning from players' previous successful runs.
- The higher imitation score after DPBM adaption shows an improvement compared to the default AI agent and the qualitative assessment shows that players perceived the DPBM agent to replicate their decision-making. However, there is still some variability since a proportion of the players did not perceive this.
- The quantitative result shows the encounter with the DPBM agent acquired higher median scores in miniPXI metrics compared to the one with the static behavior.

8 Limitation and Future Works

Despite the findings of this study, several notable limitations and areas for future improvement exist. One significant limitation is the incomplete replication of the player's complete builds, such as Cast, Call, Boons, and Hammer upgrades. This exclusion stemmed from the system complexity and the time constraints of this research. The absence of these key gameplay elements led to feedback from the community indicating lower interest in playing the mod. In addition, the lack of these features introduced bias in the research findings. Specifically, although the DPBM agent could replicate players' actions, it did not produce the same effects as the player. For example, a player with a Zeus boon equipped on dash action would release a lightning bolt while dashing, a feature that the original dash did not possess. Consequently, Dark Zagreus might use dash aggressively without causing any damage, leading to an unfair match where the DDA agent is at a disadvantage. Even if enemy parameters like health and speed were adjusted, players still found the fight relatively easy with their full builds according to the qualitative responses. Future work should aim to address these limitations by implementing the full range of player builds, including Cast, Call, Boons, and Hammer upgrades, to provide a more comprehensive assessment of the DPBM approach.

Additionally, another limitation identified in this study is the DPBM's inability to capture players' attack charge time and frequency. Charge time is crucial for players using chargeable weapons, as less proficient players may take longer to target enemies or fully charge their attacks. Similarly, attack frequency is vital; less experienced players might attack aggressively even when enemies are at a far distance. I also noticed that certain weapons require DPBM to capture frequency accurately to use those weapons effectively. This can be a direction to explore to refine the DPBM approach, ensuring it can accurately replicate a wider range of player behaviors and strategies.

Moreover, the study period might have been too brief to determine if players would continue to feel challenged and excited after multiple encounters with Dark Zagreus. This short timeframe could introduce bias, as initial encounters with a new type of boss often generate heightened excitement. Conducting a long-term study could provide more robust data on player engagement over time and yield valuable insights, making it a promising area for future research.

Within this research, an intriguing question was generated while reviewing the result: What contributes to the feeling of "fighting yourself"? Several participants rated the imitation score high even when the DDA agents were in default mode and had not yet adapted to the player's playstyle. This suggests that the perception of imitation is influenced by more than just the DPBM opponent's performance. The appearance of an opponent that resembles the player can also result in a high imitation score, regardless of the behavior. This opens up an interesting direction for research in opponent imitation perception.

By addressing these limitations, future studies can provide a more robust and accurate evaluation of the DPBM's potential to create a more engaging and challenging gaming experience in dynamic difficulty adjustment.

9 Conclusion

This research explored the integration of dynamic difficulty adjustment (DDA) via Deep Player Behavior Modeling (DPBM) into the game environment of Hades, creating a more personalized gaming experience. Despite the game system complexities and limitations encountered, the findings provide valuable insights into the potential and challenges of using deep learning models to replicate and adapt to player behaviors in real-time.

The results of the study in miniPXI metrics indicated slight improvements in metrics such as Challenge, Mastery, and Autonomy, after implementing the DPBM, suggesting that players perceived the DPBM agent as more challenging and reflective of their strategies. However, since no significant differences were found, the increases in these scores can be interpreted as noises.

A similar trend and issue are found in the imitation score, which is not enough to state a successful imitation. Luckily, the qualitative feedback still provides valuable insight regarding the DDA agent, with several participants noting that DPBM seemed to mimic their playstyles, adding a new layer of engagement and personalization to the game.

However, the study also revealed significant limitations, such as the exclusion of certain gameplay elements like Cast, Call, Boons, and Hammer upgrades, which affected the overall assessment of the DPBM's effectiveness. These limitations underscore the need for more future research to include these features and better capture player behaviors such as attack charge time and frequency.

Despite these challenges, this research contributes to the growing field of dynamic difficulty adjustment by demonstrating the feasibility of using deep learning models in a commercial game context, offering valuable directions for future work to build upon and refine these initial findings. To facilitate future research, I have made the mod and source code publicly available, enabling researchers to use it as a testbed for their studies.

In conclusion, although this research showed trivial results in enhancing player experience and creating more personalized AI opponents, it makes a further step forward to show the feasibility of applying deep learning models in more broad areas. In addition, this research also reveals the limitations and potential of the current method, inspiring researchers to explore more sophisticated and effective DDA mechanisms in video games.

References

[AM17] Dennis Ang and Alex Mitchell. "Comparing Effects of Dynamic Difficulty Adjustment Systems on Video Game Experience". In: *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. CHI PLAY '17. New York, NY, USA: Association for Computing Machinery, 2017, pp. 317–327. ISBN: 9781450348980. DOI: 10.1145/3116595.3116623. URL: https://doi.org/10.1145/3116595.3116623.

[ASO13] Justin T. Alexander, John Sear, and Andreas Oikonomou. "An investigation of the effects of game difficulty on player enjoyment". In: *Entertainment Computing* 4.1 (2013), pp. 53–62. ISSN: 1875-9521. DOI: https://doi.org/10.1016/j.entcom.2012.09.001. URL: https://www.sciencedirect.com/science/article/pii/S1875952112000134.

[Bar96] Richard Bartle. "Hearts, clubs, diamonds, spades: Players who suit MUDs". In: (June 1996).

[Ber+19] Christopher Berner et al. *Dota 2 with Large Scale Deep Reinforcement Learning*. Dec. 2019.

[CD09] Alessandro Canossa and Anders Drachen. "Patterns of Play: Play-Personas in User-Centred Game Development". In: (Jan. 2009).

[Com] SSG Modding Community. *SSG Modding Community Github*. URL: https://github.com/SGG-Modding.

[Cor09] Valve Corporation. *Left 4 Dead 2*. Video game. 2009.

[Csi90] Mihaly Csikszentmihalyi. "Flow: The Psychology of Optimal Experience". In: Jan. 1990.

[Dem+17] Simon Demediuk et al. "Monte Carlo tree search based algorithms for dynamic difficulty adjustment". In: Aug. 2017, pp. 53–59. DOI: 10.1109/CIG.2017.8080415.

[Dem+19] Simon Demediuk et al. "Challenging AI: Evaluating the Effect of MCTS-Driven Dynamic Difficulty Adjustment on Player Enjoyment". In: *Proceedings of the Australasian Computer Science Week Multiconference*. ACSW '19. Sydney, NSW, Australia: Association for Computing Machinery, 2019. ISBN: 9781450366038. DOI: 10.1145/3290688.3290748. URL: https://doi.org/10.1145/3290688.3290748.

[dFF18] Edirlei Soares de Lima, Bruno Feijó, and Antonio L. Furtado. "Player behavior and personality modeling for interactive storytelling in games". In: *Entertainment Computing* 28 (2018), pp. 32–48. ISSN: 1875-9521. DOI: https://doi.org/10.1016/j.entcom.2018.08.003. URL: https://www.sciencedirect.com/science/article/pii/S1875952118300120.

[DP19] Soumia Dermouche and Catherine Pelachaud. "Generative Model of Agent's Behaviors in Human-Agent Interaction". In: *2019 International Conference on Multimodal Interaction*. ICMI '19. Suzhou, China: Association for Computing Machinery, 2019, pp. 375–384. ISBN: 9781450368605. DOI: 10.1145/3340555.3353758. URL: https://doi.org/10.1145/3340555.3353758.

[FJP21] Pedro Fernandes, Jonathan Jørgensen, and Niels Poldervaart. *Adapting Procedural Content Generation to Player Personas Through Evolution*. Dec. 2021.

[Fro16] FromSoftware. *Dark Souls III*. Video game. 2016.

[Fro22] FromSoftware. *Elden Ring*. Video game. 2022.

[Gam20] Supergiant Games. *Hades*. Video game. 2020.

[GS11] Sam Ganzfried and Tuomas Sandholm. "Game theory-based opponent modeling in large imperfect-information games". In: *The 10th International Conference on Autonomous Agents and Multiagent Systems - Volume 2*. AAMAS '11. Taipei, Taiwan: International Foundation for Autonomous Agents and Multiagent Systems, 2011, pp. 533–540. ISBN: 0982657161.

[Hai+22] Aqeel Haider et al. "miniPXI: Development and Validation of an Eleven-Item Measure of the Player Experience Inventory". In: *Proc. ACM Hum.-Comput. Interact.* 6.CHI PLAY (Oct. 2022). DOI: 10.1145/3549507. URL: https://doi.org/10.1145/3549507.

[HB09] Stephen Hladky and Vadim Bulitko. "An Evaluation of Models for Predicting Opponent Positions in First-Person Shooter Video Games". In: Jan. 2009, pp. 39–46. DOI: 10.1109/CIG.2008.5035619.

[Hol+14] Christoffer Holmgård et al. "Generative Agents for Player Decision Modeling in Games". In: Jan. 2014.

[Hol+21] Christoffer Holmgård et al. "Monte-Carlo Tree Search for Persona Based Player Modeling". In: *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment* 11 (June 2021), pp. 8–14. DOI: 10.1609/aiide.v11i5.12849.

[Ish+18] Ryota Ishii et al. "Monte-Carlo Tree Search Implementation of Fighting Game AIs Having Personas". In: Aug. 2018, pp. 1–8. DOI: 10.1109/CIG.2018.8490367.

[KV21] Johann Knorr and Carlos Vaz de Carvalho. "Using Dynamic Difficulty Adjustment to Improve the Experience and Train FPS Gamers". In: Oct. 2021, pp. 195–200. DOI: 10.1145/3486011.3486447.

[Lav+21] Léandre Lavoie-Hudon et al. “Cognitive Shadowing for Learning Opponents in a Strategy Game Experiment: Using Machine Learning to Counter Players’ Strategies”. In: *Extended Abstracts of the 2021 Annual Symposium on Computer-Human Interaction in Play*. CHI PLAY ’21. New York, NY, USA: Association for Computing Machinery, 2021, pp. 308–313. ISBN: 9781450383561. DOI: 10.1145/3450337.3483453. URL: https://doi.org/10.1145/3450337.3483453.

[Mod] Nexus Mods. *Nexus Mods — Hades*. URL: https://www.nexusmods.com/hades.

[Moo+22] JaeYoung Moon et al. “Diversifying dynamic difficulty adjustment agent by integrating player state models into Monte-Carlo tree search”. In: *Expert Syst. Appl.* 205.C (Nov. 2022). ISSN: 0957-4174. DOI: 10.1016/j.eswa.2022.117677. URL: https://doi.org/10.1016/j.eswa.2022.117677.

[NZ22] Samer Nashed and Shlomo Zilberstein. “A Survey of Opponent Modeling in Adversarial Domains”. In: *J. Artif. Int. Res.* 73 (May 2022). ISSN: 1076-9757. DOI: 10.1613/jair.1.12889. URL: https://doi.org/10.1613/jair.1.12889.

[Pfa+20] Johannes Pfau et al. “Bot or not? User Perceptions of Player Substitution with Deep Player Behavior Models”. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI ’20. ¡conf-loc¿, ¡city¿Honolulu¡/city¿, ¡state¿HI¡/state¿, ¡country¿USA¡/country¿, ¡/conf-loc¿: Association for Computing Machinery, 2020, pp. 1–10. ISBN: 9781450367080. DOI: 10.1145/3313831.3376223. URL: https://doi.org/10.1145/3313831.3376223.

[PM21] Siphesihle Philezwini Sithungu and Elizabeth Marie Ehlers. “A Reinforcement Learning-Based Classification Symbiont Agent for Dynamic Difficulty Balancing”. In: *Proceedings of the 2020 3rd International Conference on Computational Intelligence and Intelligent Systems*. CIIS ’20. Tokyo, Japan: Association for Computing Machinery, 2021, pp. 15–23. ISBN: 9781450388085. DOI: 10.1145/3440840.3440856. URL: https://doi.org/10.1145/3440840.3440856.

[PSM18] Johannes Pfau, Jan David Smeddinck, and Rainer Malaka. “Towards Deep Player Behavior Models in MMORPGs”. In: *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*. CHI PLAY ’18. New York, NY, USA: Association for Computing Machinery, 2018, pp. 381–392. ISBN: 9781450356244. DOI: 10.1145/3242671.3242706. URL: https://doi.org/10.1145/3242671.3242706.

[PSM19] Johannes Pfau, Jan David Smeddinck, and Rainer Malaka. “Deep Player Behavior Models: Evaluating a Novel Take on Dynamic Difficulty Adjustment”. In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI EA ’19. Glasgow, Scotland Uk:

Association for Computing Machinery, 2019, pp. 1–6. ISBN: 9781450359719. DOI: 10.1145/3290607.3312899. URL: https://doi.org/10.1145/3290607.3312899.

[PSM20] Johannes Pfau, Jan David Smeddinck, and Rainer Malaka. "Enemy Within: Long-term Motivation Effects of Deep Player Behavior Models for Dynamic Difficulty Adjustment". In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, pp. 1–10. ISBN: 9781450367080. DOI: 10.1145/3313831.3376423. URL: https://doi.org/10.1145/3313831.3376423.

[Red] Reddit. *Hades Subreddit*. URL: https://www.reddit.com/r/HadesTheGame/.

[Roo+21] Shaghayegh Roohi et al. "Predicting Game Difficulty and Engagement Using AI Players". In: *Proc. ACM Hum.-Comput. Interact.* 5.CHI PLAY (Oct. 2021). DOI: 10.1145/3474658. URL: https://doi.org/10.1145/3474658.

[SBS07] Frederik Schadd, Sander Bakkes, and Pieter Spronck. "Opponent Modeling in Real-Time Strategy Games." In: Jan. 2007, pp. 61–70.

[Sil+16] David Silver et al. "Mastering the game of Go with deep neural networks and tree search". In: *Nature* 529 (Jan. 2016), pp. 484–489. DOI: 10.1038/nature16961.

[SNC17] Mirna Paula Silva, Victor do Nascimento Silva, and Luiz Chaimowicz. "Dynamic difficulty adjustment on MOBA games". In: *Entertainment Computing* 18 (2017), pp. 103–123. ISSN: 1875-9521. DOI: https://doi.org/10.1016/j.entcom.2016.10.002. URL: https://www.sciencedirect.com/science/article/pii/S1875952116300350.

[SSC15] Mirna Silva, Victor Silva, and Luiz Chaimowicz. "Dynamic Difficulty Adjustment through an Adaptive AI". In: Nov. 2015, pp. 173–182. DOI: 10.1109/SBGames.2015.16.

[Stu24] Arrowhead Game Studios. *Helldivers 2*. Video game. 2024.

[Sut+15] Rhio Sutoyo et al. "Dynamic Difficulty Adjustment in Tower Defence". In: *Procedia Computer Science* 59 (Dec. 2015), pp. 435–444. DOI: 10.1016/j.procs.2015.07.563.

[wix13] wixico. *luann*. https://github.com/wixico/luann. 2013.

[Wu+22] Zhe Wu et al. "L2E: Learning to Exploit Your Opponent". In: *2022 International Joint Conference on Neural Networks (IJCNN)*. 2022, pp. 1–8. DOI: 10.1109/IJCNN55064.2022.9892077.

10 Appendidx A - Questionnaire Example

Page 1

Introduction

Thank you for participating in our research project titled "Facing Your Self: Dynamic Difficulty Adjustment in Hades" This study aims to explore how dynamic difficulty adjustment (DDA) game agents affect player experience. As part of this research, you will play a modified version of Hades.

Purpose of the Survey

The following questionnaire is designed to gather your thoughts and feelings about the final boss during the game. Your feedback will help us understand how well the game AI's difficulty adjustments match your skill level and enhance your overall gaming experience.

Anonymity and Voluntary Participation

Please note:

All responses will be anonymous, and no personal identifying details will be collected. Although some questions ask you to upload data, this data is anonymous and does not include any personal information. You are not required to upload the data, it is optional. Also, you are free to extract the text file to review its contents before uploading. You are free to stop at any point during the survey. No answers will be recorded unless the entire survey is submitted at the end.

Consent

Before proceeding, please read the statements below and tick the final box to confirm your consent to participate in this project.

- I confirm that the research project "Facing Your Self: Dynamic Difficulty Adjustment in Hades" has been explained to me.
- I am over 18 years of age.
- I have had the opportunity to ask questions about the project and have had these answered satisfactorily.
- I have had enough time to consider whether to participate.

- I understand that my participation in this research is voluntary. I may withdraw from the study at any time without providing a reason, and any personal data already collected from me will be erased if I withdraw.
- I consent to allow the fully anonymized data to be used in future publications and other scholarly means of disseminating the findings from the research project.
- I understand that the data acquired will be securely stored by researchers, but appropriately anonymized data may be made available to others for research purposes in the future.
- I understand that the University may publish appropriately anonymized data in suitable data repositories for verification purposes and to make it accessible to researchers and other research users.

[ ] I confirm I have read and understood the above statement and agree to participate this research.

Page 2

Thank you for your time and participation. Your input is invaluable to our research. The process will take around 1 hour since you are required to play 2 runs of Hades and answer the questions in 2 questionnaires. Please follow the steps below to complete the questionnaires:

Install the Mod

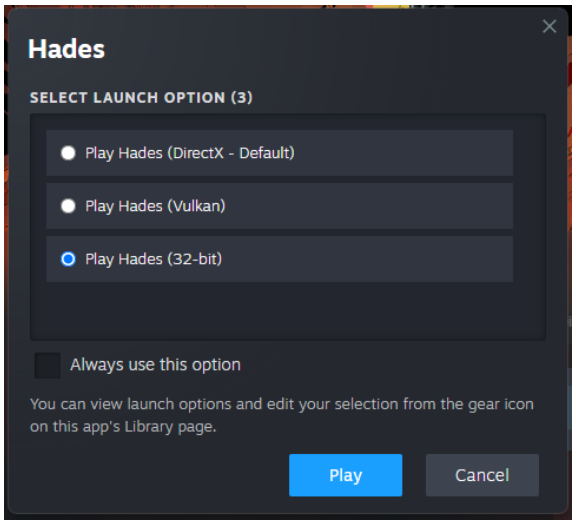
Assume you have Hades already. Download the mod from here. Follow the instructions to install the mod for the game.

Please disable all other mods for the duration of this study, so it stays comparable. Of course you are welcome to try this mod in combination with other mods after this study.

Run Hades on x86(32 bit) version (Optional)

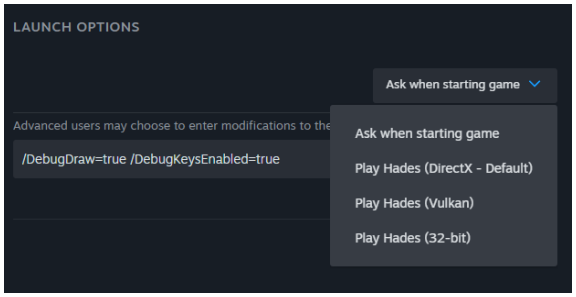
If you would like to share your play record with us, please run the game in the x86 version, which enables the export record feature. To enable the x86 version, follow these instructions:

When you start the game on Steam, there should be several options. x86 is the last option.



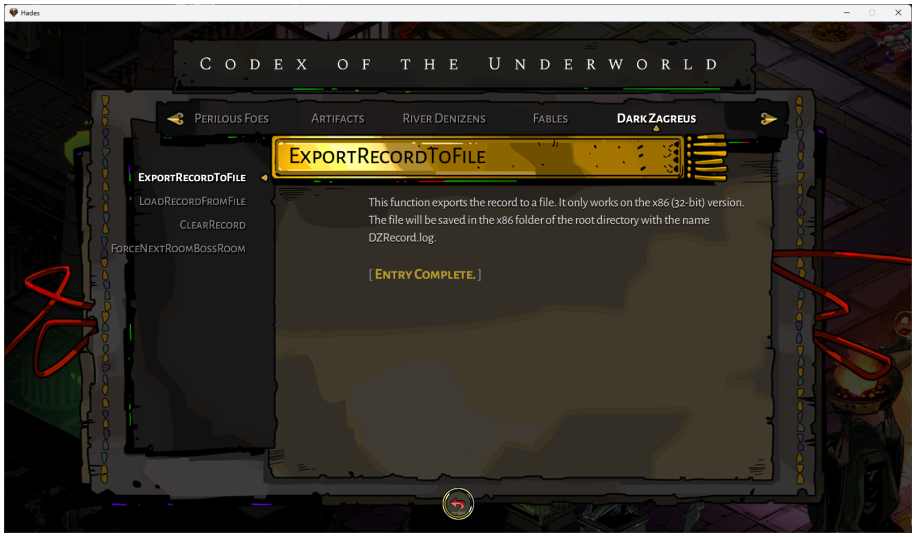
If you don't see it. You can also find it by right click the game on Steam $->$Properties $->$ General $->$ Launch Options.

Note that it is optional, you are not required to share your play record. However, it will help us a lot for future research.



Verification

To ensure the mod is successfully installed, go to the House of Hades and press C or the corresponding button to open the Codex Menu. If there is a section titled "Dark Zagreus", the mod has been successfully installed.



If you encounter any issues or have any questions during this study, please feel free to contact me. Email: h.lin2@students.uu.nl. Discord: huien

You can also find me in SSG(Hades) Modding Community Discord server by tagging @huiun.

Press the next button to start the process.

Page 3

First Playthrough

Next, start a full playthrough of the modified version of Hades. During this run, you will encounter Dark Zagreus at the final boss level. In this initial encounter, Dark Zagreus will be equipped with the basic sword weapon and his actions will be driven by random probabilities. Your goal is to complete the run by defeating Dark Zagreus. You are free to choose any weapon except for the following weapons:

- Aspect of Hera (Bow3)
- Aspect of Zeus (Shield3)
- Aspect of Beowulf (Shield4)

These weapons are not supported due to technical issues.

Please avoid building a Cast build for Zagreus, as Dark Zagreus currently cannot utilize Cast due to technical limitations.

Please disable any heat difficulty level modifiers before starting the run to keep it comparable.

First Questionnaire

After completing the first run, you will need to fill out a questionnaire based on your experience during this playthrough.

Important: If you die during the first run, please repeat it, until you have successfully defeated Dark Zagreus for the first time.

Answer all the questions thoroughly. Press the next button to proceed to the questionnaires.

Page 4

First Playthrough

Please rate the following statements based on your experience with Dark Zagreus during the game.

- 1Q1: Which weapon were you using? (Weapon name)
- 1Q2: Which variant of weapon were you using? (Variant index)
- 1Q3: Dark Zagreus was not too easy and not too hard to play against. (1 to 5)
- 1Q4: I felt free to fight against to Dark Zagreus in my own way (1 to 5)
- 1Q5: I was curious to see how Dark Zagreus would react to my actions. (1 to 5)
- 1Q6: I was fully focused on playing against Dark Zagreus. (1 to 5)
- 1Q7: I felt that fighting Dark Zagreus kept me engaged. (1 to 5)
- 1Q8: Overcoming Dark Zagreus gave me a sense of mastery over the game. (1 to 5)
- 1Q9: Playing against Dark Zagreus felt like a valuable experience. (1 to 5)
- 1Q10: I had a lot of fun playing against Dark Zagreus. (1 to 5)
- 1Q11: Fighting Dark Zagreus felt like combating myself. (1 to 5)
- 1Q12: Please summarize your experience of fighting against Dark Zagreus in a few words. (Open-ended)

1Q13: Upload Your First Playthrough Record (Optional)
You can find your playthrough record by using the export command in the Codex Menu, under the Dark Zagreus section. The export file will be placed at location. Uploading this file can help us analyze and improve the AI's performance. Note that this data is anonymous and does not include any personal information. You may review the contents of the file before uploading.

Page 5

Second Playthrough

After filling out the first questionnaire, begin a second full playthrough. If possible, please complete both playthroughs on the same day, ideally right after each other.

This time, when you reach the final boss level, you will face Dark Zagreus again. However, in this encounter, Dark Zagreus will be a little different. You are free to choose any weapon.

Important: If you die before you make it to the final boss, please repeat the second run. Should you succeed or fail against Dark Zagreus now, you can proceed with the questionnaire.

Second Questionnaire

After completing the second run, fill out the second questionnaire based on your experience during this playthrough. Again, answer all the questions thoroughly and honestly to help us gather comprehensive data.

Page 6

Second Playthrough

Please rate the following statements based on your experience with Dark Zagreus during the game.

- 2Q1: Which weapon were you using? (Weapon name)
- 2Q2: Which variant of weapon were you using? (Variant index)
- 2Q3: Dark Zagreus was not too easy and not too hard to play against. (1 to 5)
- 2Q4: I felt free to fight against to Dark Zagreus in my own way (1 to 5)
- 2Q5: I was curious to see how Dark Zagreus would react to my actions. (1 to 5)
- 2Q6: I was fully focused on playing against Dark Zagreus. (1 to 5)
- 2Q7: I felt that fighting Dark Zagreus kept me engaged. (1 to 5)
- 2Q8: Overcoming Dark Zagreus gave me a sense of mastery over the game. (1 to 5)
- 2Q9: Playing against Dark Zagreus felt like a valuable experience. (1 to 5)
- 2Q10: I had a lot of fun playing against Dark Zagreus. (1 to 5)
- 2Q11: Fighting Dark Zagreus felt like combating myself. (1 to 5)
- 2Q12: Please summarize your experience of fighting against Dark Zagreus in a few words. (Open-ended)

2Q13: Upload Your Second Playthrough Record (Optional)
You can find your playthrough record by using the export command in the Codex Menu, under the Dark Zagreus section. The export file will be placed at location. Uploading this file can help us analyze and improve the AI's performance. Note that this data is anonymous and does not include any personal information. You may review the contents of the file before uploading.

Page 7

Thank you for your time and participation. Your input is invaluable to our research. The following questions are optional, and you may choose whether or not to answer them.

Would you like to share suggestions or thoughts on improving the boss behavior or the mod in general? Please feel free to provide any additional feedback below. (Open-ended)

11 Appendidx B - Raw Data

ID	1Q1	1Q2	1Q3	1Q4	1Q5	1Q6	1Q7	1Q8	1Q9	1Q10	1Q11
0	Bow	2	2	5	5	5	5	4	5	5	1
1	Sword	1	5	4	4	5	4	2	3	4	4
2	Bow	1	1	3	3	2	2	1	2	1	4
3	Bow	2	1	4	4	5	2	1	2	3	2
4	Shield	4	5	5	5	5	5	5	5	5	4
5	Gun	2	2	5	5	2	4	3	4	4	4
6	Sword	4	3	5	5	5	5	5	5	5	3
7	Gun	1	2	4	4	3	2	4	4	4	1
8	Bow	2	4	3	5	3	4	3	4	3	3
9	Gun	3	1	5	5	4	1	1	2	3	4
10	Bow	2	4	4	5	5	3	1	2	4	4
11	Sword	3	2	5	5	4	4	4	5	5	5
12	Bow	1	2	5	5	5	4	4	5	3	2
13	Gun	3	3	5	5	4	3	2	4	5	2
14	Fist	3	4	3	5	5	4	3	2	4	4
15	Sword	1	4	5	5	5	5	4	5	5	5
16	Fist	1	4	1	1	5	2	1	3	1	1
17	Spear	2	3	4	5	5	3	3	2	3	3
18	Bow	2	4	5	4	5	3	4	3	4	1
19	Gun	3	4	4	4	5	5	5	3	4	1

Table 9. *The questionnaire result of the first playthrough.*

ID	1Q12
0	I am impressed with this BOSS, and I find its skill set unfamiliar and novel.
1	It was some trouble as i couldnt really anticipate most of the attacks due to the small stature of zagreus and the fact that most of the time when i attacked dark zagreus and zagreus were covered with splashes of colours from the boons
2	Really basic for a final boss
3	This is a very very cool concept! I think that the fight overall on 0 heat was extremely easy (took around 30 damage total). The fight was heavily in my favor though since I had the bow and chill effects
4	
5	high damage "speedrunner" eris rail build killed him too fast for me to see too many base sword mechanics, though excited to see what happens in the second run
6	Greate work for huain
7	Awkward
8	because i was using a long range weapon and he sarted with the sword it was a bit difficult to create distance but otherise a fun fight
9	i kept dashing pass him and shooting him with the empowered shot and he couldn't keep up
10	Interesting, but seems unbalanced as a boss due to screen size and lack of telegraphing attacks which is common across all bosses. Option to export playthrough data wasn't an option for some reason.

Table 10. *The open-ended responses of the first playthrough.*

ID	1Q12
11	I felt it was slightly easy (maybe because my build was too strong), but I did feel like it mirrored my playstyle! I'm usually very up-close and brawly and Dark Zagreus definitely played the same!
12	I don't usually use the sword, and the fight itself was kind of easy, just took a bit to learn the attack pattern.
13	I did not use that sword and it's my first time to see him. I don't feel like he's acting like me anyways. But it was fun to fight against.
14	I did it without a build so I took long and I lost
15	challenging and gives a good insight on your own playstyle
16	zagreus has player style animations meaning they aren't telegraphed. dodging was near impossible and required cheese trats (dashing away with zeus on my dash)
17	
18	I'm new to the game which made getting to the final boss a challenge, but once I did it felt like i was playing against a much more skilled player
19	

Table 11. *The open-ended responses of the first playthrough.*

ID	2Q1	2Q2	2Q3	2Q4	2Q5	2Q6	2Q7	2Q8	2Q9	2Q10	2Q11
0	Spear	2	1	5	4	4	5	5	5	5	5
1	Shield	2	2	5	4	2	3	3	4	4	4
2	Shield	1	1	3	3	2	2	1	2	2	1
3	Gun	3	5	3	4	4	2	1	2	2	2
4	Sword	1	5	5	5	5	5	5	5	5	4
5	Fist	1	1	5	5	5	5	5	5	5	1
6	Sword	2	2	5	5	5	5	5	5	5	5
7	Bow	2	3	4	4	4	4	2	3	2	2
8	Bow	2	4	5	5	5	4	4	3	4	5
9	Spear	4	4	5	5	5	5	4	4	5	4
10	Gun	2	2	5	5	5	2	1	2	2	2
11	Bow	4	4	5	5	5	5	4	4	5	5
12	Sword	2	5	5	5	4	4	4	4	5	5
13	Spear	2	4	5	5	4	4	3	4	5	4
14	Fist	3	5	5	5	3	1	3	3	4	1
15	Sword	2	5	5	5	5	5	5	5	5	5
16	Bow	2	3	4	4	5	4	3	3	4	3
17	Spear	2	4	5	5	5	5	5	5	4	5
18	Shield	2	2	4	5	2	3	2	2	4	5
19	Gun	3	2	4	5	5	5	3	5	4	5

Table 12. *The questionnaire result of the second playthrough.*

ID	2Q12
0	I UESD A BROKEN BUILD lol.
1	it was attacking in a different pattern though most of my previous boons were of critical in nature which with a shield in hand never really touched me so it was fairly easy though i did enjoyed the mod and am curious to explore more
2	When I play bow I like to stay close and dash-strike repeatedly while dodging, DZ was doing long charges most of the time , and threw some specials between attacks, which I never do. It was a very easy fight, especially using shield.
3	Maybe it was just the aspect of Chiron, but again the damage was way too low and since his aspect was only at level one, it was extremely easy. I will try again at a higher heat to see if this helps combat how quickly DZ gets rolled. A second phase would be very cool too but I am sure that it is difficult to implement lol
4	

Table 13. *The open-ended responses of the second playthrough.*

ID	2Q12
5	i couldn't seem to find the playthrough records sorry :(my overall thoughts are: DZ needs more health. a lot of hades boss fights boil down to me dash spamming to avoid 50-75% of damage hits and bursting bosses down before they can kill me, and since dark zagreus has much less hp than hades it feels a lot easier -because of the above point i couldn't really see the adaptive difficultyAI at work, kind of killed him within 1020 seconds. maybe give dark zagreus a subset of the boons from our previous run?
6	SO MUCH FUN
7	
8	Definitely felt like i was fighting against my own scummy tactics
9	he kept reloading with out shooting around 70% of the time but when he was working it was very scary and fun
10	Good to see that it seemed like only the loadout and not the play style that was copied. Zag's loadout doesn't seem tuned to fight against, but is obviously very fun to fight with. Having Hades' voice lines, nothing for either good or bad Zag felt a bit immersion breaking but otherwise was an interesting fight.
11	
12	I used the sword this time since I seldom use it anyway, but I knew he'd use the bow against me. Pretty fun, though I do feel I definitely bullied him a bit with Poseidon special haha

Table 14. *The open-ended responses of the second playthrough.*

ID	2Q12
13	I saw him taking advantage of special attack like me. It's kind of interesting. But still I feel like it's too easy to defeat him. If I want, he won't even have any chance to hurt me. Maybe he got more chance if he has the same boons as I had.
14	I did a run where I took Zeus and Poseidon and it was easy to finish the boss
15	a hard but valuable challenge that helped me improve the run after. which was quicker and with less deaths
16	relatively easy fight, but all the odds were in my favour. i had a speed boon, Poseidon boons kept it at a distance so despite it constantly running at me i could sometimes shoot it and then hit it a lot with my special from the aspect of chiron with doom. it sometimes got a hit on me which did not feel fair due to no signposting due to PC animations, but overall an easy fight.
17	It was really fun to feel like you are fighting yourself
18	It felt quite sad to see my poor skills reflected in dark zagreus... Having had more experience since the first round, I found the second fight way easier both because I was better at the game but also because the enemy was slower to respond and I had way more time to do my attacks.
19	

Table 15. *The open-ended responses of the second playthrough.*

12 Appendidx C - Statistical Results

show all result graphs, expecting 3 - 5 pages

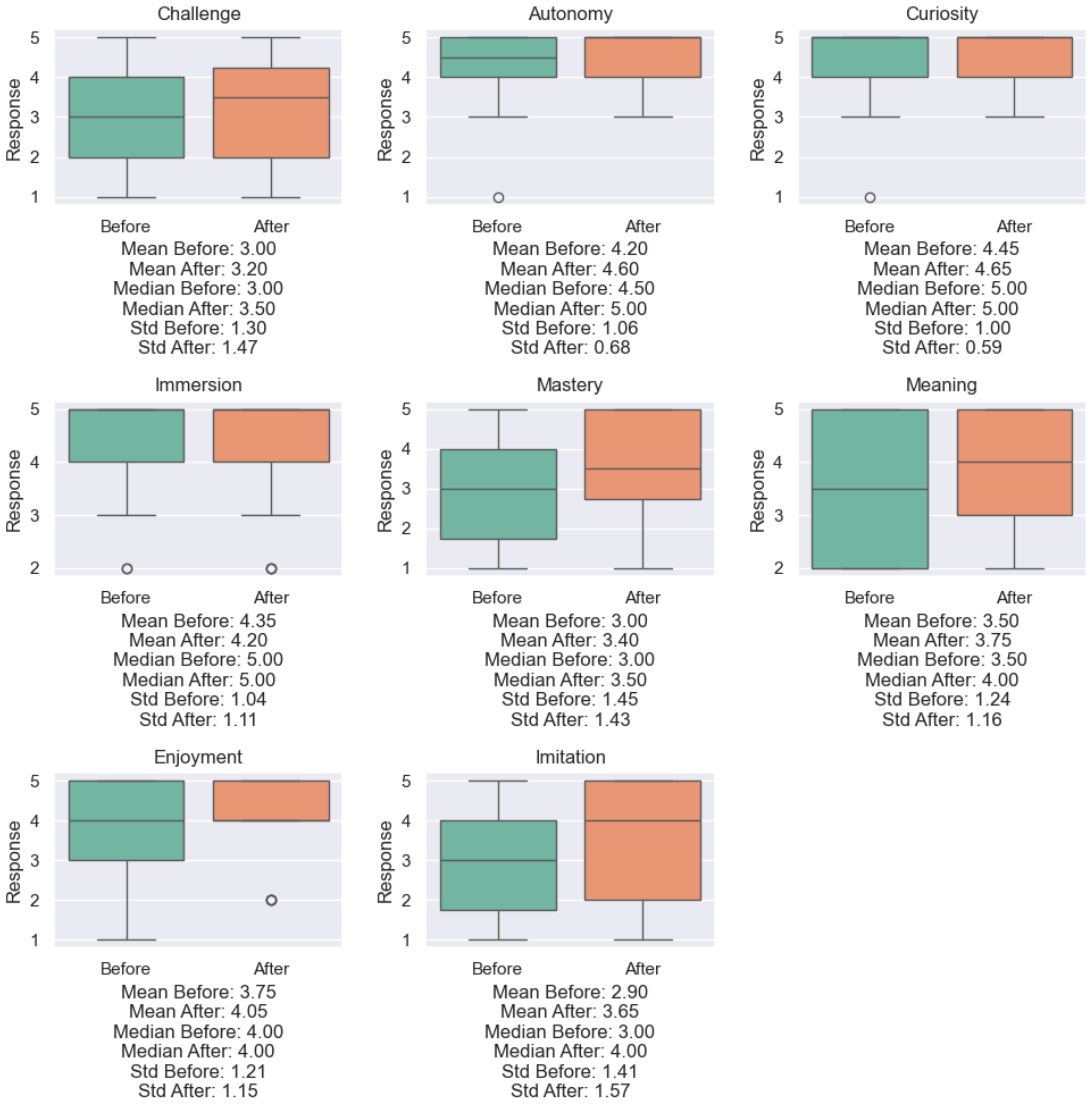


Figure 17. *The result of miniPXI metrics.*

Weapon	Variant Index	Dash Toward	Attack	Special Attack	Dash Away	Advanced Attack	Reload	Accuracy
Bow	3	818	664	519	706	0	0	0.58
Bow	2	524	689	596	741	0	0	0.60
Gun	3	332	1969	0	414	0	340	0.73
Gun	3	480	723	0	541	0	518	0.63
Spear	2	409	404	211	590	5	0	0.47
Spear	4	432	870	13	449	70	0	0.50
Spear	2	575	1205	618	544	0	0	0.54
Spear	2	348	1935	319	370	8	0	0.72
Sword	4	1634	874	17	1811	0	0	0.43
Sword	3	593	881	361	496	0	0	0.49
Sword	2	850	465	558	903	0	0	0.40

Table 16. *The voluntarily uploaded records.*

13 Appendidx D - Github Repository and Mod Link

Dark Zagreus Github Repository

https://github.com/willake/hades-dark-zagreus-mod

Prediction Accuracy Analysis Code

https://github.com/willake/hades-dark-zagreus-model-validator

Dark Zagreus Nexus Mod Page

https://www.nexusmods.com/hades/mods/191

14 Appendidx E - Ethical Approvals

		Yes	No
P7	Do you intend to be alone with a research participant or have to take sole responsibility for the participants at any point during your research activity?		V

If P7 is no continue with P8, otherwise:

As you will be alone with or solely responsible for vulnerable participants (yes to P7) a fuller ethical review is required. You may also need a Certificate of Conduct (Dutch: VOG) from the government. Please add more detail here:

		Yes	No
P8	Does your project involve participants with whom you have, or are likely to have, a working or professional relationship: for instance, staff or students of the university, professional colleagues, or clients?		V

If the answer to P8 is yes, please answer P9, otherwise, continue with PC1.

		Yes	No
P9	Is it made clear to potential participants that not participating will in no way impact them (e.g. it will not directly impact their grade in a class)?		V

If the answer to P9 is yes, then continue with PC1, otherwise:

As participants may think that not participating may harm them (yes to P8 and no to P9), participation may no longer be voluntary. Hence, a fuller ethical review is required. Please provide more information here:

Consent Procedures		**Yes**	**No**	**Not applicable**
PC1	Do you have set procedures that you will use for obtaining *informed* consent from all participants, including (where appropriate) parental consent for children or consent from legally authorized representatives? (See suggestions for information sheets and consent forms on the website.)	V		
PC2	Will you tell participants that their participation is voluntary?	V		
PC3	Will you obtain explicit consent for participation?	V		
PC4	Will you obtain explicit consent for any sensor readings, eye tracking, photos, audio, and/or video recordings?	V		
PC5	Will you tell participants that they may withdraw from the research at any time and for any reason?	V		
PC6	Will you give potential participants time to consider participation?	V		
PC7	Will you provide participants with an opportunity to ask questions about the research before consenting to take part (e.g. by providing your contact details)?	V		

If the answer to PC1-PC7 is yes, then continue with PC8, otherwise:

Given your responses to the informed consent questions (a no on any of PC1-PC7), a fuller ethical review is required. Please provide more information regarding the questions that are causing this here:

		Yes	**No**
PC8	Does your project involve concealment or deliberate misleading of participants?		V

If the answer to PC8 no, continue with Section 2, otherwise:

 As you plan to use concealment or misleading (yes to PC8), and this may impact participants' rights to informed consent, a fuller ethical review is required. Please provide more information on the concealment/misleading here:

Section 2. Data protection, handling, and storage

The General Data Protection Regulation imposes several obligations for the use of **personal data** (defined as any information relating to an identified or identifiable living person) or including the use of personal data in research.

		Yes	No
D1	Are you gathering or using personal data (defined as any information relating to an identified or identifiable living person)?		V

If the answer to D1 is yes, please answer the following questions; otherwise, continue with Section 3.

High-Risk Data

		Yes	No
DR1	Will you process personal data that would jeopardize the physical health or safety of individuals in the event of a personal data breach?		V
DR2	Will you combine, compare, or match personal data obtained from multiple sources, in a way that exceeds the reasonable expectations of the people whose data it is?		V
DR3	Will you use any personal data of children or vulnerable individuals for marketing, profiling, automated decision-making, or to offer online services to them?		V

DR4	Will you profile individuals on a large scale?		V
DR5	Will you systematically monitor individuals in a publicly accessible area on a large scale (or use the data of such monitoring)?		V
DR6	Will you use special category personal data, criminal offense personal data, or other sensitive personal data on a large scale?		V
DR7	Will you determine an individual's access to a product, service, opportunity, or benefit based on an automated decision or special category personal data?		V
DR8	Will you systematically and extensively monitor or profile individuals, with significant effects on them?		V
DR9	Will you use innovative technology to process sensitive personal data?		V

If the answer to DR1-DR9 is no, continue with DM1, otherwise:

 **As high-risk data processing seems involved (yes to any of DR1-DR9), a fuller privacy assessment is required.** Please provide more information on the DR1-DR9 questions with a yes here:

Data Minimization

		Yes	No
DM1	Will you collect only personal data that is strictly necessary for the research?	V	

If you answered yes to DM1 continue with DM4, otherwise:

		Yes	No
DM2	Will you only collect not strictly necessary personal data because it is (1) technically unfeasible not to collect it when collecting necessary data, or (2) needed as a source of necessary data?		
DM3	Will you (1) extract any necessary data as soon as possible from the		

	collected not strictly necessary data and (2) delete the not strictly necessary data immediately after any required extraction?		
DM4	Will you anonymize the data wherever possible?	V	
DM5	Will you pseudonymize the data if you are not able to anonymize it, replacing personal details with an identifier, and keeping the key separate from the data set?	V	

If the answer to any of DM2-DM5 is no, see warning below, otherwise continue with DC1.

As you do not seem to minimize data collection (no to any of DM2-DM5), a fuller privacy assessment is required. Please provide more information on the DM2-DM5 questions with a no here:

Using Collaborators or Contractors that Process Personal Data Securely

		Yes	**No**
DC1	Will any organization external to Utrecht University be involved in processing personal data (e.g. for transcription, data analysis, data storage)?		V

If the answer to DC1 is yes, please complete DC2 otherwise continue with DI1.

		Yes	**No**
DC2	Will this involve data that is not anonymized?		

If the answer to DC2 is yes, please complete DC3-DC5, otherwise continue with DI1.

		Yes	**No**	**Not Applicable**
DC3	Are they capable of securely handling data?			
DC4	Has been drawn up in a structured and generally agreed manner who is responsible for what concerning data in the collaboration?			

DC5	Is a written contract covering this data processing in place for any organization which is not another university in a joint research project?			

If the answer to any of DC3-DC5 is no, see warning below, otherwise continue with DI1.

 **As you do not seem to have appropriate processes in place for sharing data with collaborators or contractors (no to any of DC3-DC5), a fuller privacy assessment is required.** Please provide more information on the DC3-DC5 questions with a no here:

International Personal Data Transfers

		Yes	**No**
DI1	Will any personal data be transferred to another country (including to research collaborators in a joint project)?		V

If the answer to DI1 is yes, please complete DI2, otherwise continue with DF1.

		Yes	**No**
DI2	Do all countries involved in this have an adequate data protection regime?		

If the answer to DI2 is no, please complete DI3, otherwise continue with DF1.

		Yes	**No**
DI3	Is a legal agreement in place?		

If the answer to DI2 and DI3 is no, see warning below, otherwise, continue with DF1.

 **As you do not seem to have appropriate safeguards in place for international data transfers (no to DI2 and DI3), a fuller privacy assessment is required.** Please provide more information on intended international data transfers here:

Fair Usage of Personal Data to Recruit Participants

		Yes	No
DF1	Is personal data used to recruit participants?		V

If the answer to DF1 is yes please answer DF2-DF4, otherwise continue with DP1

		Yes	No	N/A
DF2	Have potential participants provided this personal data voluntarily to be contacted about the research or is the data publicly available?			
DF3	If contact details have been provided by a third party, would participants expect their details to be passed on to the university and to be used in this way?			
DF4	If contact details have been gathered for a purpose other than research, would participants expect their details to be used in this way?			

If the answers to DF2-DF4 are yes or N/A continue with DP1, otherwise:

As there seem to be issues with your use of personal data for recruitment (no to one or more of DF2-DF4), a fuller privacy assessment is required. Please provide more information on the intended use of personal data for recruitment here:

Participants' data rights and privacy information

		Yes	No	Not Applicable
DP1	Will participants be provided with privacy information? (Recommended is to use as part of the information sheet: For details of our legal basis for using personal data and the rights you have over your data please see the University's privacy information at www.uu.nl/en/organisation/privacy.)		V	
DP2	Will participants be aware of what their data is used for?	V		
DP3	Can participants request that their personal data be deleted?			V
		Yes	**No**	**Not Applicable**
DP4	Can participants request that their personal data be rectified (in case it is incorrect)?[24]			V
DP5	Can participants request access to their personal data?[24]			
DP6	Can participants request that personal data processing is restricted?			
DP7	Will participants be subjected to automated decision-making based on their personal data with an impact on them beyond the research study to which they consented?			
DP8	Will participants be aware of how long their data is being kept for, who it is being shared with, and any safeguards that apply in case of international sharing?			
DP9	If data is provided by a third party, are people whose data is in the data set provided with (1) the privacy information and (2) what categories of data you will use?			

If the answer to DP1-DP6, DP8, DP9 is yes and DP7 is no, continue with DE1, otherwise:

As there seem to be issues with the data rights of your participants or the provision of privacy information (no to one or more of DP1-DP6, DP8, DP9, or yes to DP7), a fuller privacy assessment is required. Please provide more detail regarding data rights and/or privacy information here:

Using data you have not gathered directly from participants

		Yes	No

DE1	Will you use any personal data that you have not gathered directly from participants (such as data from an existing data set, data gathered by a third party, data scraped from the internet)?		V

If the answer to DE1 is no please continue with DS1.

		Yes	**No**
DE2	Will you use an existing dataset in your research?		

If the answer to DE2 is yes please answer DE3-DE5, otherwise, continue with DE6.

		Yes	**No**
DE3	Do you have permission to do so from the owners of the dataset?		
DE4	Have the people whose data is in the data set consented to their data being used by other researchers and/or for purposes other than that for which that data set was gathered?		
DE5	Are there any contractual conditions attached to working with or storing the data from DE2?		

		Yes	**No**
DE6	Does your project require access to personal data about participants from other parties (e.g., teachers, employers), databanks, or files?		

If the answer to DE6 is yes please answer DE7-DE8, otherwise, continue with DE9.

		Yes	**No**
DE7	Do you have a process in place to gain informed consent from these participants?		
DE8	Are there any contractual conditions attached to working with or storing the data from DE5?		

		Yes	**No**
DE9	Does the project involve collecting personal data from websites or social media (e.g., Facebook, Twitter, Reddit)?		

 **As there may be issues with the use of existing data (no to DE3, DE4, DE7 or yes to DE9), a fuller privacy assessment is required.** Please provide more detail regarding the use of existing data here:

Ethics and Privacy Quick Scan (version: 31 July 2023)

Section 1. Research projects involving human participants

		Yes	No
P1	Does your project involve human participants? This includes for example use of observation, (online) surveys, interviews, tests, focus groups, and workshops where human participants provide information or data to inform the research. If you are only using existing data sets or publicly available data (e.g. from Twitter, Reddit) without directly recruiting participants, please answer no.	V	

If no, continue with Section 2; if yes, fill in the following questions.

Recruitment

		Yes	No
P2	Does your project involve participants younger than 18 years of age?		V
P3	Does your project involve participants with learning or communication difficulties of a severity that may impact their ability to provide informed consent?		V
P4	Is your project likely to involve participants engaging in illegal activities?		V
P5	Does your project involve patients?		V
P6	Does your project involve participants belonging to a vulnerable group, other than those listed above?		V

If the answer to all of P2-P6 is no, continue with P8.

As you are dealing with vulnerable participants (yes to one (or more) of P2-P6) a fuller ethical review is required. Please add more detail on your participants here:

Secure data storage

		Yes	No
DS1	Will any data be stored (temporarily or permanently) anywhere other than on password-protected University authorized computers or servers?		V
If the answer to DS1 is yes, please answer DS2, otherwise, continue with DS4.			
		Yes	**No**
DS2	Does this only involve data stored temporarily during a session with participants (e.g. data stored on a video/audio recorder/sensing device), which is immediately transferred (directly or with the use of an encrypted and password-protected data-carrier (such as a USB stick)) to a password-protected University authorized computer or server, and deleted from the data capture and data-carrier device immediately after transfer?		

If the answer to DS2 is yes, continue with DS4, otherwise answer DS3.

		Yes	No
DS3	Does this only involve data stored with a collaborator or contractor?		
DS4	Excluding (1) any international data transfers mentioned above and (2) any sharing of data with collaborators and contractors, will any personal data be stored, collected, or accessed from outside the EU?		V

If the answer to DS2 and DS3 is no, or the answer to DS4 is yes, see the warning below, otherwise continue with Section 3.

As there may be issues with secure data storage (no to DS2 and DS3, or yes to DS4), a fuller privacy assessment is required. Please provide more detail regarding data storage here:

Section 3: Research that may cause harm

Research may harm participants, researchers, the university, or society. This includes when technology has dual-use, and you investigate an innocent use, but your results could be used by others in a harmful way. If you are unsure regarding possible harm to the university or society, please discuss your concerns with the Research Support Office.

		Yes	No
H1	Does your project give rise to a realistic risk to the national security of any country?		V
H2	Does your project give rise to a realistic risk of aiding human rights abuses in any country?		V
H3	Does your project (and its data) give rise to a realistic risk of damaging the University's reputation? (E.g., bad press coverage, public protest.)		V
H4	Does your project (and in particular its data) give rise to an increased risk of attack (cyber- or otherwise) against the University? (E.g., from pressure groups.)		V
H5	Is the data likely to contain material that is indecent, offensive, defamatory, threatening, discriminatory, or extremist?		V
H6	Does your project give rise to a realistic risk of harm to the researchers?		V
H7	Is there a realistic risk of any participant experiencing physical or psychological harm or discomfort?		V
H8	Is there a realistic risk of any participant experiencing a detriment to their interests as a result of participation?		V
H9	Is there a realistic risk of other types of negative externalities?		V

If the answer to H1-H9 is no continue with Section 4, otherwise:

 **As you replied yes to one (or more) of H1-H9, a fuller ethical review is required.** Please provide more detail here on the potential harm, and how you will minimize risk and impact:

Section 4: Conflicts of interest

		Yes	No
C1	Is there any potential conflict of interest (e.g. between research funder and researchers or participants and researchers) that may potentially affect the research outcome or the dissemination of research findings?		V
C2	Is there a direct hierarchical relationship between researchers and participants?		V

If the answer to C1-C2 is yes, continue with Section 5, otherwise:

 **As you replied yes to C1 or C2, a fuller ethical review is required.** Please provide more information regarding possible conflicts of interest and how you mitigate them here:

Section 5: Your information

This last section collects data about you and your project so that we can register that you completed the Ethics and Privacy Quick Scan, sent you (and your supervisor) the summary of what you filled out, and follow up where a fuller ethics review and/or privacy assessment is needed. For details of our legal basis for using personal data and the rights you have over your data please see the University's privac HYPERLINK "http://www.uu.nl/en/organisation/privacy"y HYPERLINK "http://www.uu.nl/en/organisation/privacy" info HYPERLINK "http://www.uu.nl/en/organisation/privacy"r HYPERLINK "http://www.uu.nl/en/organisation/privacy"mation. Please see the guidance on the ICS Ethics and Privacy website on what happens on submission.

Z0. Which is your main department?

● Information and Computing Science

○ Freudenthal Institute

○ Pharmacy

○ Other, namely:

Z1. Your full name:

Hui En Lin

Z2. Your email address:

h.lin2@students.uu.nl

Z3. In what context will you conduct this research?

○ 1. As a student on a course with course coordinator:

○ 2. As a student for my bachelor thesis, supervised by:

● 3. As a student for my master thesis, supervised by: Dr. J. Pfau and Dr S.C.J. Bakkes

○ 4. As a PhD student, supervised by:

○ 5. As an independent researcher (e.g. research fellow, assistant/associate/full professor)

In case the answer to Z3 is 2:

Z4. Bachelor programme for which you are doing the thesis:

○ Artificial Intelligence (Kunstmatige Intelligentie)

○ Computing Science (Informatica)

○ Information Science (Informatiekunde)

○ Other:

In case the answer to Z3 is 3:

Z5. Master programme for which you are doing the thesis:

○ Applied Data Science

○ Artificial Intelligence

○ Business Informatics

○ Computing Science

○ Data Science

● Game and Media Technology

○ Human-Computer Interaction

○ Other:

In case the answer to Z3 is 1, 2, 3, or 4:

Z6. Email of the course coordinator or supervisor (so that we can inform them that you filled this out and provide them with a summary):

j.pfau@uu.nl

In case the answer to Z3 is 2 or 3:

Z7. Email of the moderator (as provided by the coordinator of your thesis project):

gmt-ethics@uu.nl (Julian Frommel)

Z8. Title of the research project/study for which you filled out this Quick Scan:

Facing Your Self: Dynamic Difficulty Adjustment in Hades via Deep Player Behavior Modeling

Z9. Summary of what you intend to investigate and how you will investigate this (200 words max):

In our study, we aim to investigate player experience with a dynamic difficulty adjustment agent, which adapts its behavior according to the player's progression. A local deep-learning model is trained using the player's gameplay records; however, this data is not transmitted back to us. Instead, we will develop a mod for an existing game and release it within the modding community. Alongside the mod, we will provide a questionnaire for players to voluntarily complete. Players are free to install and engage with our mod, and participation in the questionnaire is entirely optional. It is made clear to players that they have the choice to fill out the questionnaire provided. Importantly, no personal information is collected during the research process. The questionnaire will include queries based on the Intrinsic Motivation Inventory (IMI), aimed at evaluating player enjoyment and engagement.

In case the answer to Z3 is 2 or 3:

		Yes	No	Not Applicable
Z10.	In case you encountered warnings in the survey, does your supervisor already have ethical approval for a research line that fully covers your project?			

In case the answer to Z10 is yes:

Z11. Provide details on the ethical approval (e.g. ethical approval number):
