Dynamic Difficulty Adjustment for Digital Games using Reinforcement Learning

Adrian Mercieca
Institute of Information and Communication Technology
MCAST
Paola, Malta
adrian.mercieca.f31954@mcast.edu.mt

Mr Ivan Briffa
Institute of Information and Communication Technology
MCAST
Paola, Malta
ivan.briffa@mcast.edu.mt

***Abstract*—This research gauges the benefits and downsides of including overarching Dynamic Difficulty Adjustment (DDA) procedures in digital games, such that the difficulty adapts automatically to the user's performance. An Artificial Intelligence Model was developed to implement and manage DDA in 'Ravenfield'. Users were asked to try both versions, through which important data was extracted. This data showed that a system which is well implemented is beneficial to players, increasing levels of enjoyment and engagement. Systems which fail to stay anonymous can instead compel the user to feel cheated. This review highlights the need for further research on the subject, as to find an ideal way of implementing a system which is beneficial for users across the board.**

***Index Terms*—Dynamic Difficulty Adjustment, Artificial Intelligence, Digital Games**

I. Introduction

Over the past few decades, video games have become a more and more common type of entertainment. The difficulty in developing games that are both entertaining and available to a variety of players, regardless of skill level, is growing as the video game business expands. The use of Dynamic Difficulty Adjustment (DDA) in games, which entails automatically modifying the difficulty level of the game in response to the player's performance, is one remedy for skill level-related inaccessibility. DDA has the ability to increase players' enjoyment of games by giving them a personalized gaming experience that is catered to their skills.

Recently, excellent Language Models that have taken the globe by storm have helped Artificial Intelligence (AI) further establish itself as a potent technology. Real-time player performance analysis using AI algorithms enables the game to dynamically modify the challenge level to the player's proficiency level. DDA is successful in boosting player pleasure and engagement while also making games more accessible to players of all skill levels.

The purpose of this study is to investigate how DDA might be implemented in video games using AI. This study specifically looks at a DDA model that is implemented through reinforcement learning and assesses how the DDA processes affect player engagement, satisfaction, and accessibility.

Overall, the study aims to add to the expanding body of knowledge about the use of AI in video games and to offer new information about how to set up a DDA system that is both effective and efficient.

A. Research Onion

1) Research Philosophy: This research study explored the idea that social phenomena are inherently subjective and context-dependent, and can be understood through interpretation and analysis of meaning. Therefore the first layer chosen was 'Interpretivism' as this promoted understanding the subjective experience and perspective of the players. Qualitative research was used to uncover the complex and nuanced meanings that people attach to their experiences and actions.

2) Research Approach: An Inductive approach was chosen as this research's purpose is to uncover themes and patterns from the data collected. Since this topic is still relatively unexplored, this exploratory approach seemed ideal to further the topic's understanding, by arriving at a Hypothesis through the data itself.

3) Research Strategy: Data was gathered using an experiment. This experiment compared the normal version of a 'Ravenfield' match with another 'Ravenfield' match which was assisted by DDA, and managed through an AI model. The results of these two scenarios were compiled and compared with one another to produce important insight on the topic at hand. https://www.overleaf.com/project/64612545dbd7809c9a52f4e1

4) Choices: The data was gathered through the use of Mixed-Methods. Participants were asked to answer several questions, some of which were closed-ended (quantitative) and some of which were open-ended (qualitative) and promoted further discussion. All data gathered was necessary to further understand underlying factors which affect enjoyment and engagement during play.

5) Time-Horizon: Data was gathered at a single point in time, therefore the Cross-Sectional method was used. The data was however gathered from different groups of people, with varying levels of skill in the topic that was tested.

6) Techniques and Procedures: The data was gathered through post-experiment interviews which gauged the difficulty, enjoyment, and engagement during play sessions, after which a number of questions are asked that directly relate to DDA and its utility and perceived value. The sampling strategy chosen was stratified sampling. The participants were selected to represent four different sub-groups (Teenage Gamers, Teenager Non-Gamers, Adult Gamers, and Adult Non-Gamers). The researchers then analyzed the data using both content analysis and statistical analysis.

B. Researcher Positioning

The researcher has a background in digital games, both in development and in playing during leisure time. The researcher acknowledges and has experienced firsthand the emotions and detachment that can be experienced when faced with difficulty which is not in sync with the player's skill.

C. Purpose of Research

The research aims to study how DDA systems could be implemented in a way that amplifies beneficial emotions for the players without adding non-beneficial emotions such as confusion, frustration, and feeling discredited.

D. Research Questions

- Does the adjustment affect player performance?
- Do players notice when adjustments occur?
- Does adjustment significantly affect the player's enjoyment, frustration, or perception of game difficulty?

II. Literature Review

A. Game Adjustment

Game systems are iteratively modified by game creators utilizing playtesting feedback, until the game is balanced. Directed mathematical analysis can reveal deeper structures and relationships within a gaming system, even though this method cannot be automated. With the correct algorithms and methodology, these deeper structures and relationships can be changed as the game is being played [1]. Developers can utilize the "Change Blindness" phenomenon, which is the inability to identify changes when they take place during saccades, blinks, blank screens, movie cuts, and other interruptions, to make these transitions seamless [2].

B. Precautions of DDA

[1] says that it is essential to identify and comprehend the systems that make a game fun and how these can be altered to heighten this enjoyment in order to customize a gaming experience for a certain player without negatively affecting well-balanced systems. The author also emphasizes that changing the difficulty during a game could make the player feel deceived if it interferes with or undermines the basic player experience. Mechanics, Dynamics, and Aesthetics (MDA) will differ between games and genres, but they must be taken into account and taken into account when developing DDA. The various player states and game-related actions are referred to as mechanics. Dynamics is the term used to describe how a game's rewards and levels of difficulty change over time. A game's mechanics and dynamics have a significant impact on its aesthetics. The overall aesthetics of a video game are frequently influenced by user expectations based on genre standards [1]. The study also emphasized the importance of obtaining accurate measurements, such as taking damage readings during battles rather than throughout the entire playing session in order to obtain a reading that is pertinent to the combat.

C. Implementation 1

[1] developed a system that modifies negative feedback without altering the fundamental FPS genre experience using the MDA architecture as a guide. The created system manages the game's basic exploration and combat dynamics, as well as the key inventory mechanics (health, ammo, shields, and weapons), while also keeping the general cycle of activities, which in turn preserves the game's fast-paced shooter aesthetics. In order to estimate the likelihood of mortality in a specific encounter, [1] used damage sustained, health, the mean and standard deviation of current damage rates, and time. This estimate aids in determining if intervention through difficulty adjustment is necessary. [1] emphasizes how crucial it is for every DDA controller to have Adjustment Goals. The Comfort policy, which attempts to keep the players reasonably safe, is the first of the author's three separate policies that are mentioned. Instead, the Discomfort policy is designed to make players work harder by restricting resources and upping the difficulty when a player reaches a certain level of health. The author goes on to say that the Training policy initially gives the player comfort before progressively making them more uncomfortable during a level or session. The player's probability of dying has a predefined threshold of 40% in [1], and if it is exceeded, difficulty changes are made, adding 15 health points every 100 ticks.

D. Implementation 2

According to [3], DDA can aid in retaining the player's attention for a longer period of time in hypercasual endless games. The study demonstrates that while the majority of these games initially succeed in holding the player's attention, most lose their appeal after seven days of play. [3] contends that this results from boredom or irritation brought on by an imbalance between the difficulty of the game and the player's competence. The game can remain in the "Flow" state, which is when the player encounters the optimal challenge and abilities balance, by using a successful DDA system. The "core loop" is encouraged to be entered by achieving and sustaining the player's flow throughout the tutorial or introduction. The study by [3] makes use of a hypercasual infinite game that employs a Quick Progressive Difficulty (QPD) algorithm and develops an alternative version that uses DDA. Because of the way the QPD is designed, difficulty always increases when a player reaches a plateau, which can lead to dissatisfaction and worry. According to the research, dealing with these strong, unfavorable feelings through DDA may help the player stay in the "flow" and play for extended periods of time. The player starts the game with 5 lives, and after gaining 5 consecutive points without losing a life, the player can advance a level, increasing the challenge. The user is stuck at this difficulty after this level up, or harder if he or she can again accumulate 5 consecutive points without losing a life. The DDA-implemented modified version maintains the player's ability to level up in the same way but adds a requirement that allows the player to level down in order to lessen the frustration brought on by the extremely challenging gameplay. If a player loses a life following an increase in difficulty, this may mean that the difficulty may be out of whack with their current abilities, which results in a decrease in difficulty.

E. Implementation 3

[4] implemented DDA to a beloved classic called Pac-Man. The DDA is based on the use of the Monte Carlo Tree Search (MCTS), which is an improved version of the Monte Carlo Simulation algorithm. The Ghosts, which are controlled by MCTS, engage in battles against a computer-simulated Pac-Man that follows a predetermined strategy. At each decision point in the maze, the MCTS simulation constructs a search tree encompassing all the permissible moves for the Ghosts. The available legal moves include Right, Left, Up, and Down, while any moves that would result in the Ghosts or Pac-Man hitting a wall or colliding with each other are eliminated. [4] expected this system to outperform other DDA implementations in three fronts. Firstly, this adjustment is centred around the intelligence level of opponents, ensuring that players engage with adversaries possessing similar capabilities, thereby promoting fairness. Also, it provides a means of ongoing and adaptable adjustment of the game's challenge. This is achieved by modifying the simulation time of the MCTS algorithm, which is employed to generate opponent artificial intelligence. Lastly, DDA based on MCTS is implemented through extensive computation, thereby minimizing the need for domain-specific knowledge and human involvement.

F. Implementation 4

[5] explored and analyzed mDDA (multiplayer Dynamic Difficulty Adjustment) which is a gameplay element found in competitive multiplayer video games. Its purpose is to diminish the gap in challenge experienced by all players by modifying the performance of specific individuals. A notable example of this can be observed in the racing game Mario Kart 7. In this game, when players ranked lower in the race collect a weapon box, their likelihood of receiving a more effective weapon is heightened. This adjustment grants lower-performing players an increased opportunity to enhance their ranking. A total of 180 games were chosen for review using Metacritic [9], which assigns normalized scores to games based on approved game review publications and websites. The selection of games for formal review was based on their Metacritic scores, ensuring a variety of game quality levels. The initial step involved examining game genres on Metacritic, and the three genres with the highest proportion of competitive multiplayer gameplay modes were identified as First-Person Shooters (FPS), Racing, and Fighting. Subsequently, 60 games with competitive multiplayer modes were selected from each of these three genres. To encompass a diverse range of games, we opted for 30 games in each genre that had a Metacritic rating higher than 75, as well as 30 games with ratings ranging from 50 to 74 within each genre. Games with ratings below 50 were excluded due to limited information available from other sources, making them challenging to investigate effectively.

G. Implementation 5

In [6], a series of experiments is detailed, wherein Artificial Neural Networks (ANN) are trained offline. These ANNs serve as embedded game agents within a shooting game, controlling the behavior of non-player characters (NPCs). The training datasets for the ANNs are constructed based on the gameplay of three distinct player skill levels: expert, medium, and beginner. Each skill level's training dataset is then used to train a corresponding ANN. To determine the optimal number of neurons in the hidden layer and the appropriate duration for training, a three-fold cross-validation method is employed. Additionally, the study includes a comparison between the ANN approach and two traditional game AI techniques: finite state machine (FSM) and computer random controlled methods.

H. Experiments

[1] initiated the experiment by giving subjects a description of the game and its controls, followed by at least 15 minutes of playtime. Performance data was stored for later revision and subjects were then requested to fill out an evaluation form. The experiment was single-blind, with half of the games adjusted and the other half control [1]. Adversely, [3] had participants first try one version of the game, and then try the other. This allowed the research to distinguish between the results obtained from each participant for both versions. Scores and frustration evaluations were gathered from each participant after each play session. [4] gathered data in relation to using different simulation times. For each simulation time, 250 gameplays were conducted, and from these an average win rate was calculated. [5] recorded three key elements from each of the games reviewed: the activation trigger, the affected game rules and the scope of the effects. These were generalized and split further into specific components. A framework was developed to classify and recognize instances of mDDA by identifying the shared components and potential attributes associated with each component.

I. Results

[1] initiated the experiment by giving subjects a description of the game and its controls, followed by at least 15 minutes of playtime. Performance data was stored for later revision and subjects were then requested to fill out an evaluation form. The experiment was single-blind, with half of the games adjusted and the other half control [1]. Adversely, [3] had participants first try one version of the game, and then try the other. This allowed the research to distinguish between the results obtained from each participant for both versions. Scores and frustration evaluations were gathered from each participant after each play session. [4] gathered data in relation

to using different simulation times. For each simulation time, 250 gameplays were conducted, and from these an average win rate was calculated. [5] recorded three key elements from each of the games reviewed: the activation trigger, the affected game rules and the scope of the effects. These were generalized and split further into specific components. A framework was developed to classify and recognize instances of mDDA by identifying the shared components and potential attributes associated with each component. In [6], the implementation and analysis of game AI agents using ANN were described. The training data, recorded from players of different skill levels (expert, medium, and beginner) playing against computer random-controlled NPCs, were utilized to identify valuable feature vectors as inputs for the neural network. Through three-fold cross-validation, the optimal number of neurons in the hidden layer and the appropriate training duration were determined. Subsequently, the trained neural network was integrated into the game as the artificial intelligence for the NPCs. Finally, the fuzzy control theory was employed to facilitate dynamic difficulty adjustment in the game. In contrast to the traditional game AI implemented using FSM, the utilization of neural networks for decision-making in [6] brought it closer to human thinking patterns and enhanced the level of intelligent reasoning. To enhance the game's AI and introduce a wider range of behavior patterns, a trained neural network was integrated into a real-time game. This integration aimed to enhance the overall gameplay experience and increase the enjoyment factor by offering more diverse and engaging gameplay options.

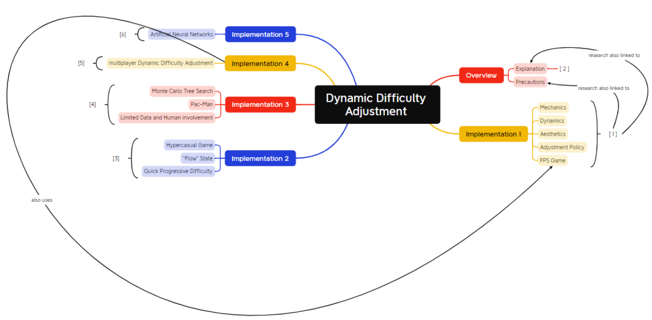

Fig. 1. Literature Map

III. Research Methodology

A. Game and Equipment Used

The video game "Ravenfield" was modified for this study to add DDA. Changes in opponent accuracy, dependent on the team's remaining unit lives and the balance of flag control, caused this revision. If there was a balance, the level of difficulty was perfect. A change was required to rebalance the teams' remaining unit lives (tickets) and/or flag balance when there was an imbalance on either side. An ASUS ROG STRIX G15 with the following specifications was used during the entire research project:

- Central Processing Unit: Intel I7-11850H
- Memory / RAM: 16GB
- Graphical Processing Unit: Nvidia RTX 2060 (Laptop)

B. Dynamic Difficulty Adjustment Models

The "stablebaselines3" Python package was used to build the AI Model. This package makes it easier to create AI models with different settings. It was found that a Multi Input Policy setting would be perfect for this investigation to handle the inevitably complicated observations. Data that the model identifies with an action during training is referred to as an observation. The initial observation listed the number of tickets left on the blue team's and red team's teams, the balance of the flags, and the level of difficulty of the current set. After initializing a class that handles logging and saving during training, the model was configured to store its best version after every 1000 steps. This made it possible to compare the effectiveness and pace of improvement of several models in great detail. The models were now be loaded and tested by playing a loaded model and returning the total rewards, which could then be compared to the output of other models. The initial model underwent five days of nonstop training and 265000 steps.

C. Reward System

This study featured a number of requirements that, when met, gave incentives to the AI Model. The first requirement relates to the difference in team tickets (unit balance), wherein the Model was given a reward with a value of 1 if the difference was 5 tickets or fewer. When the Model met another requirement—having the flag balances of the two teams within 10% of one another—another prize of the same value was obtained concurrently. When the first two requirements were met at the same time, the Model received another reward with a value of 1.

D. Gauging User Feedback

A four-phase experiment that was developed and carried out yielded useful information. To collect as much supporting information as feasible within the constraints of time, a number of questions were created for each phase. Nineteen individuals were asked to participate in the experiment and accepted. The participants were split into two groups, as was done by [8], with the only difference being the sequence in which the steps were completed. Phases 2 and 3 were completed by the first group before phase 2, whereas phase 3 was completed by the second group before phase 2. This was done to reduce bias from questions following a game, where a person's responses would have been strongly influenced by their adaptibility which is measured in phase 1 through self-assessment. There were several questions written for each stage. Primary data that was gathered came mostly from these questions.

E. The Experiment's Process

1) Phase 1:

- How often do you play digital games? (1-5)
- How would you rate your own skill in playing First Person Shooter Games? (1-5)
- How would you rate your adaptability to new game experiences? (1-5)
- Which of these do prefer the game to be: Easy, Challenging, or Extremely Challenging?
- Have you ever played 'Ravenfield'? If yes, give an estimation of total play-time.

Each participant received a personalized explanation of the game's rules. in particular, the loadout choice, the manual reloading, and the rag dolling effect. The controls were then briefly explained to the participants. Then it was stated that after answering several pre-play questions, players would be required to play a single match before receiving additional instructions. The Phase 1 questions, which prompted self-assessment on experiment-related skills and assessed the participants' expectations for the result of their future play session, were then given to the participants to complete.

2) Phase 2:

- How challenging was the match? (1-5)
- How engaged were you during play? (1-5)
- How much would you say you enjoyed it? (1-5)
- How well-balanced was the game's difficulty? (1-5)
- Did you notice anything strange during the match?

The game 'Ravenfield' in its original form must be used to play a match, according to the instructions given to the participants. They were free to employ any loadout they pleased, and they were told to play until the match was over, whether they won or lost. Participants were required to respond to questions on the current Phase after the gaming session. These inquiries measured the participant's perceptions of the questions' enjoyment, engagement, and difficulty.

3) Phase 3:

- How challenging was the match? (1-5)
- How engaged were you during play? (1-5)
- How much would you say you enjoyed it? (1-5)
- How well-balanced was the game's difficulty? (1-5)
- Did you notice anything strange during the match?

With one significant difference, this Phase functioned exactly like Phase 2. In this stage, the players engaged in a match using Dynamic Difficulty Adjustment with an altered version of the game "Ravenfield." Evaluation was completed following the conclusion of the game, just as it did in Phase 2.

4) Phase 4: The research and the notion of dynamic difficulty adjustment were ultimately explained to the participants and any final thoughts were noted down.

IV. Results

A. Perceived Challenge

The participants provided feedback after both matches played which can be compared in order to identify differences in gameplay such as the perceived difficulty. After phase 2, most participants agreed that the difficulty was easy to medium as shown in Figure 2. Contrastingly, when playing the math with DDA in phase 3, most participants agreed that the challenge aspect had risen and was now medium to hard, as shown in Figure 3. Due to alternating matches during the experiment stage, the result also factors in participant adaptability.

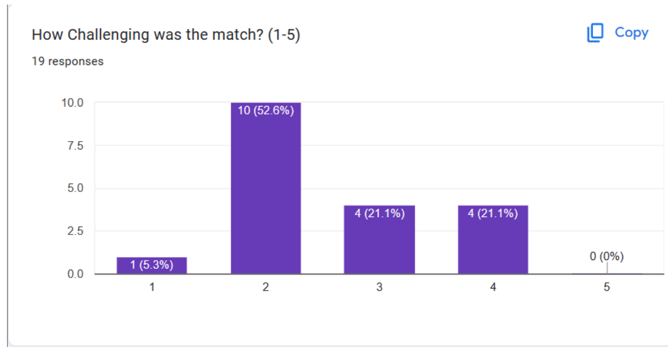

Fig. 2. Difficulty: Phase 2 Results

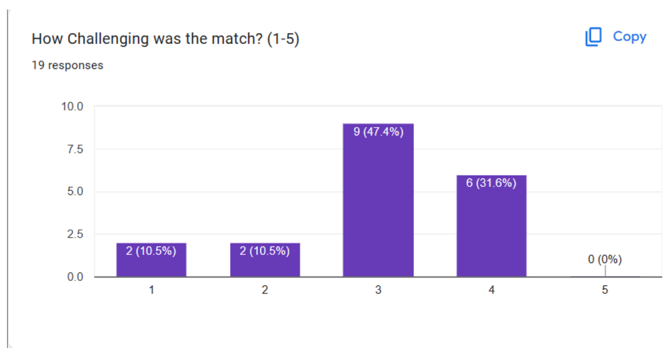

Fig. 3. Difficulty: Phase 3 Results

B. Engagement

During phase 2, most participants were highly engaged when facing the normal game difficulty, however a few outliers felt out of sync with the game as shown in Figure 4. No huge changes were noted in the data collected after phase 3, as only a minor shift in the lower scores of engagement was noted, as shown in Figure 5.

C. Enjoyment

As with engagement, both gameplays yielded very similar results. In phase 2, most participants highly enjoyed the match, as seen in Figure 6. However, in phase 3 a slight increase in enjoyment was again noted by the participants, as shown in Figure 7.

D. Team Balance

Lastly, participants were asked to indicate how balanced the battle felt during both individual matches. With regards to phase 2's gameplay, Figure 8 shows that the majority of

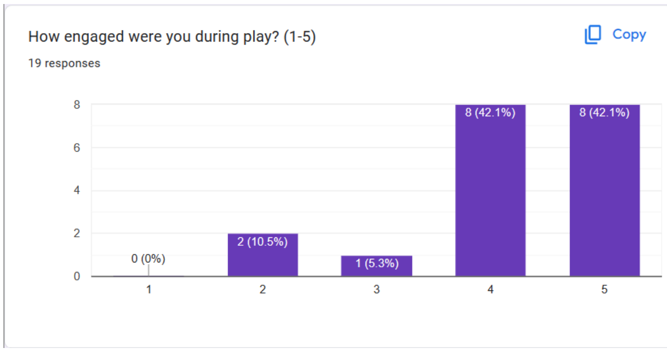

Fig. 4. Engagement: Phase 2 Results

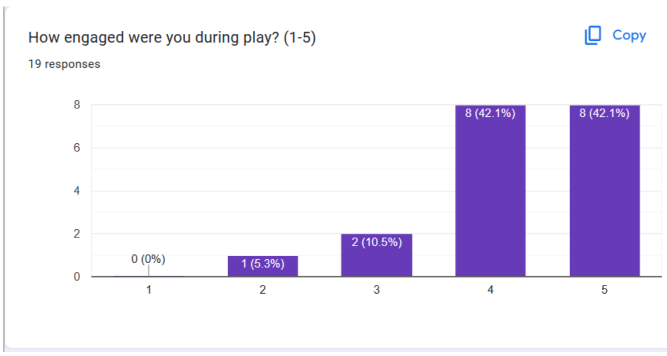

Fig. 5. Engagement: Phase 3 Results

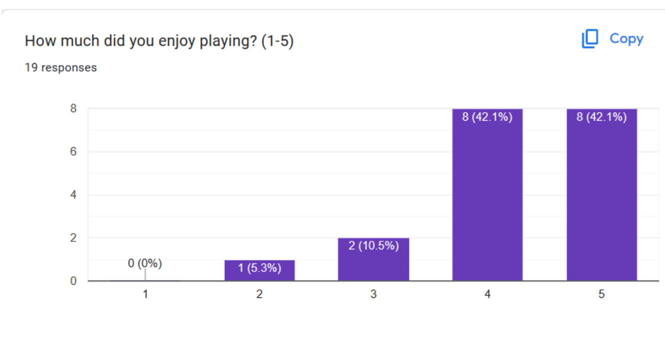

Fig. 6. Enjoyment: Phase 2 Results

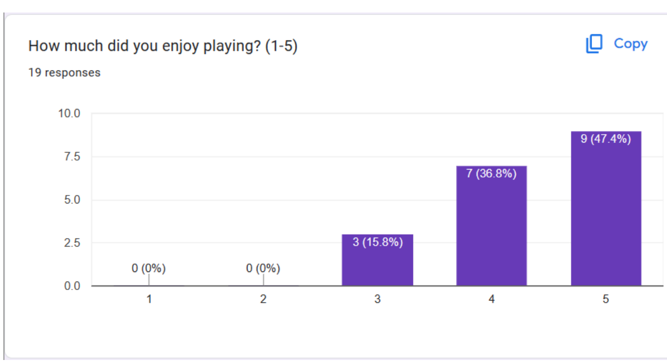

Fig. 7. Enjoyment: Phase 3 Results

the participants felt a lack in balance between the two teams, Whereas, during phase 3, teams were noted to have improved balance as can be seen in Figure 9 and were closer in final scores.

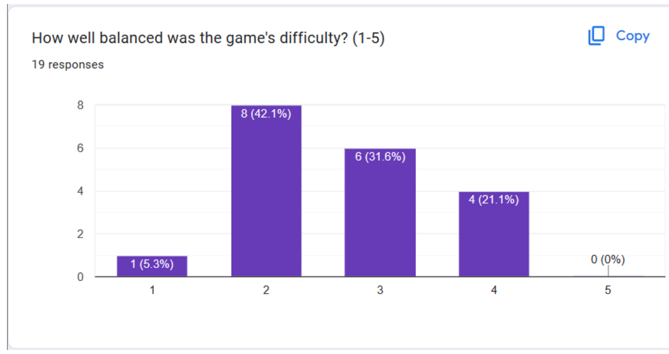

Fig. 8. Balance: Phase 2 Results

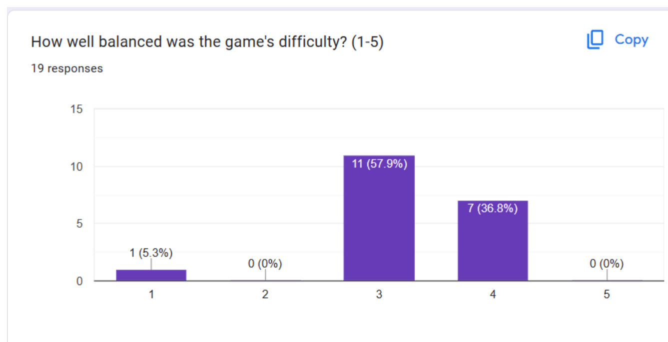

Fig. 9. Balance: Phase 3 Results

E. *Change Blindness*

The DDA was, for the most part, unnoticed. This could have been due to the large number of factors and NPC's which were in the battle: 25 on each side. Figure 10 shows how only 2 participants managed to identify the shift in difficulty as strange when compared to the previously played battle, most credited their adaptability or accused their NPC teammates of under-performing this time around.

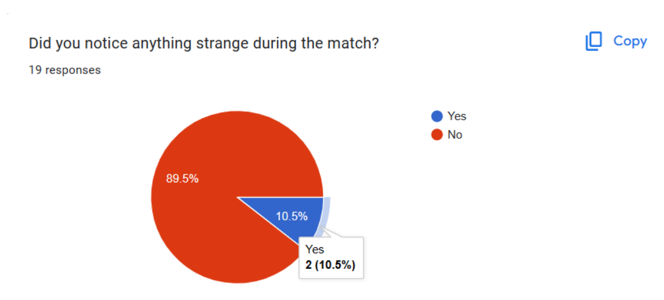

Fig. 10. Suspicion Identification

V. Conclusion

The results of this research are positive and encourage further and more in depth research on the use and ideal implementation of DDA through the use of AI. Participants felt more engaged, felt more enjoyment and felt better overall balance with DDA, however, during phase 4 some remarked this adjustment could cause a player to feel cheated, and that it may blind the user to any progress he/she makes, due to the game adjusting to match the newly attained skill level. The methodology could have been improved by reducing the number of NPC's present during the matches played. This would have decreased variability, getting more accurate results from the AI and providing a clearer understanding of whether the number of NPC's affect Change Blindness' effectiveness. Another shortcoming was the lack of resources available, as a controller setup would have been ideal to some of the participants, therefore they forced to adapt to a new game while adapting to input channels which are uncommonly used by them. Further research should be carried out on implementing such AI powered DDA systems in Multiplayer games were the player is matched with NPC's. Having such a difficulty adjustment could lead to more engagement from all players, but it might also aggravate players into feeling cheated or unskilled. This merits further investigation as it could open the doors to more competitive play in the future.

References

[1] R. Hunicke, "The Case for Dynamic Difficulty Adjustment in Games," pp. 429–433, Jun. 2005, doi: 10.1145/1178477.1178573.

[2] D. J. Simons, "Current Approaches to Change Blindness," vol. 7, no. 1–3, pp. 1–15, 2000, doi: 10.1080/135062800394658.

[3] Z. Yang and B. Sun, "Hyper-Casual Endless Game Based Dynamic Difficulty Adjustment System For Players Replay Ability," 2020, pp. 860–866.

[4] J.-Y. Wang and Y.-R. Tseng, "Dynamic difficulty adjustment by fuzzy rules using in a neural network controlled game," 2013, pp. 277–281, doi: 10.1109/ICNC.2013.6817985.

[5] A. Baldwin, D. Johnson, P. Wyeth, and P. Sweetser, "A framework of Dynamic Difficulty Adjustment in competitive multiplayer video games," 2013, pp. 16–19, doi: 10.1109/IGIC.2013.6659150.

[6] Y. Hao, S. He, J. Wang, X. Liu, jiajian Yang, and W. Huang, "Dynamic Difficulty Adjustment of Game AI by MCTS for the game Pac-Man," 2010, vol. 8, pp. 3918–3922, doi: 10.1109/ICNC.2010.5584761.