



Article

Exploring Dynamic Difficulty Adjustment Methods for Video Games

Nicholas Fisher * and Arun K. Kulshreshth *

School of Computing and Informatics, University of Louisiana at Lafayette, Lafayette, LA 70503, USA

* Correspondence: c00413598@louisiana.edu (N.F.); arunkul@louisiana.edu (A.K.K.)

Abstract: Maintaining player engagement is pivotal for video game success, yet achieving the optimal difficulty level that adapts to diverse player skills remains a significant challenge. Initial difficulty settings in games often fail to accommodate the evolving abilities of players, necessitating adaptive difficulty mechanisms to keep the gaming experience engaging. This study introduces a custom first-person-shooter (FPS) game to explore Dynamic Difficulty Adjustment (DDA) techniques, leveraging both performance metrics and emotional responses gathered from physiological sensors. Through a within-subjects experiment involving casual and experienced gamers, we scrutinized the effects of various DDA methods on player performance and self-reported game perceptions. Contrary to expectations, our research did not identify a singular, most effective DDA strategy. Instead, findings suggest a complex landscape where no one approach—be it performance-based, emotion-based, or a hybrid—demonstrably surpasses static difficulty settings in enhancing player engagement or game experience. Noteworthy is the data's alignment with Flow Theory, suggesting potential for the Emotion DDA technique to foster engagement by matching challenges to player skill levels. However, the overall modest impact of DDA on performance metrics and emotional responses highlights the intricate challenge of designing adaptive difficulty that resonates with both the mechanical and emotional facets of gameplay. Our investigation contributes to the broader dialogue on adaptive game design, emphasizing the need for further research to refine DDA approaches. By advancing our understanding and methodologies, especially in emotion recognition, we aim to develop more sophisticated DDA strategies. These strategies aspire to dynamically align game challenges with individual player states, making games more accessible, engaging, and enjoyable for a wider audience.

**Citation:** Fisher, N.; Kulshreshth, A.K.Exploring Dynamic Difficulty Adjustment Methods for Video Games. *Virtual Worlds* **2024**, *3*, 230–255. <https://doi.org/10.3390/virtualworlds3020012>

Academic Editor: Kostas Karpouzis

Received: 29 February 2024

Revised: 17 April 2024

Accepted: 31 May 2024

Published: 7 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: dynamic difficulty adjustment; video games; game design; gameplay metrics; physiological sensors; electroencephalography (EEG); user experience; user study

1. Introduction

Player engagement is pivotal for the success of a video game. If a game lacks intuitive gameplay and engaging mechanics, players might be tempted to switch to a different activity. This is particularly true for casual and non-gamers, who possess a diverse range of other recreational options. If these players are not immediately engaged and enjoy the game, they may set the game aside and never give it another chance. The game's difficulty significantly influences engagement; if it is too challenging, players may give up, whereas if it is too easy, they might lose interest. However, designing difficulty tiers that accommodate the diverse skill set of potential players is challenging. Factors like player skill, gaming experience, and familiarity with a game's genre are important in determining the optimal difficulty. A potential solution is to adjust the game's difficulty based on the player's performance and emotional state, gauged by using physiological sensors. Dynamically modifying game difficulty, considering players' affective states and in-game performance can enhance playtime and satisfaction. Such adaptability is important for enticing newcomers to video gaming and maintaining interest among those who do not typically engage with games.

Game designers traditionally offer fixed difficulty options, such as easy, medium, and hard (i.e., difficult), from which players can choose. However, two primary issues emerge with this approach. Firstly, these fixed settings might not cater to the full spectrum of player capabilities. Even the easiest difficulty level may prove to be too challenging for a non-gamer. For a player who may not have any experience with the controller or input system, any fixed difficulty level would prove too challenging for that user. Likewise, it is difficult to provide an appropriate challenge for experienced gamers. The toughest difficulties may not provide an engaging experience because even the highest level of difficulty available may not provide enough of a challenge. Secondly, these static difficulties do not accommodate fluctuating moods. Typically, games permit difficulty selection at the beginning of the game or specific checkpoints. Yet, a player's mood can vary frequently, influencing their gameplay preferences at any given moment. The player may get frustrated by continually failing and may just want to move past this specific point. Or, after a long day of work, they may just want to relax and not challenge themselves as much as they usually would. Addressing players' affective states when determining difficulty could notably bolster engagement.

Dynamic Difficulty Adjustment (DDA), which can consider the players' skill, emotional state, and game context, promises a more captivating experience than traditional games with fixed difficulties. This approach could be particularly beneficial for players at the skill extremes, potentially fostering continuous engagement and heightened immersion—an important element in game enjoyment [1]. While DDA is not new, it traditionally emphasizes performance metrics [2]. However, recent studies have begun exploring affective-based DDA as well [3–5].

In this study, we introduce and evaluate five distinct difficulty modification approaches. We crafted a custom game featuring seven difficulty tiers and assessed it using these five methods: fixed, user-selected, performance-based, emotion-based, and a hybrid of performance and emotion. Performance metrics were derived from in-game parameters, while emotional data were sourced from physiological sensors. We conducted a comprehensive user study with 31 participants, comparing the methods based on performance and self-reported measures, such as stress and engagement.

2. Related Work

DDA serves as an important mechanism in video game design, addressing the challenges of maintaining player engagement across diverse skill levels and emotional states. Unlike static difficulties that cannot adapt to individual player needs, DDA can align game challenges with player's abilities and mood in real-time, ensuring a continuous state of engagement and immersion. These approaches, leveraging real-time adjustments based on performance metrics or physiological responses, represent a significant advancement over traditional methods. The work of Bicalho et al. [6] identified player classification as a key challenge in DDA. By facilitating a more personalized experience, DDA not only broadens the appeal of video games but also offers an avenue to transform the short-lived interest of non-gamers and casual players into lasting engagement.

Many approaches to performance-based or emotion-based DDA derive inspiration from Csíkszentmihályi's Flow Theory [7]. This theory suggests that a player's enjoyment and experience peak when tasked with an appropriately challenging endeavor. Describing a state of intense concentration or complete absorption in the ongoing activity, the phenomenon is often dubbed as being "in the zone". A flow state is achieved when there is an equilibrium between the player's skill level and the game's difficulty as illustrated in Figure 1. By adopting real-time difficulty adjustments as opposed to static settings, games can potentially heighten immersion. Building on the subject of player experience, Bowman and Tamborini [8] delved into the ramifications of stress and boredom on task demand in gaming contexts. Additionally, Jenett et al. [1] explored the interplay between immersion in games and corresponding affective measures. Lemmons [9] found that matching player

skill to difficulty led to higher levels of flow as well as better performance, higher arousal, and more enjoyment.

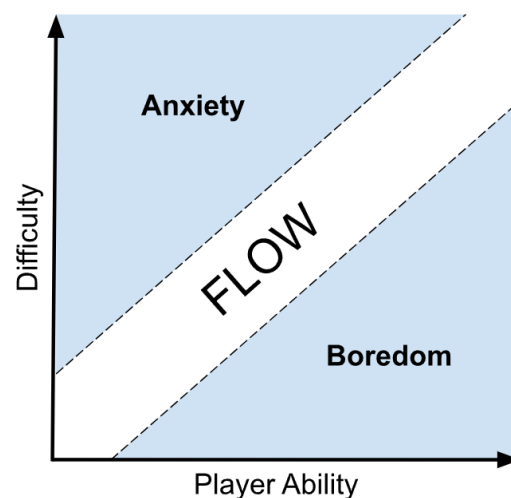


Figure 1. Flow state is achieved by presenting an appropriate challenge for the players skill level. If the game is too difficult, the player is more stressed and has higher anxiety. If the game is too easy, the player is less engaged and becomes bored.

Many studies have explored DDA and its benefits [2]. Specifically, performance-based DDA emerges as a prevalent alternative to conventional static game difficulties. This approach employs in-game performance metrics, such as accuracy, damage dealt, and total score, to adjust difficulty dynamically. Research on performance-based DDA spans genres, from first-person shooters [10–12] to Massively-Multiplayer Online (MMO) games [13] and Multiplayer Online Battle Arena (MOBA) titles [14]. Romero-Mendez et al. [15] utilized deep learning to classify players' skill levels and applied performance-based DDA, significantly enhancing gameplay engagement and immersion in an arcade shooter. In the context of enhancing player engagement in VR exercise games, Huber et al. [16] proposed using procedurally generated environments alongside DDA to dynamically tailor game levels to players' physical capabilities. Francillette et al. [17] developed and tested a model for performance-based DDA within a platformer environment, assessing difficulty through static danger zones and enemy movement patterns.

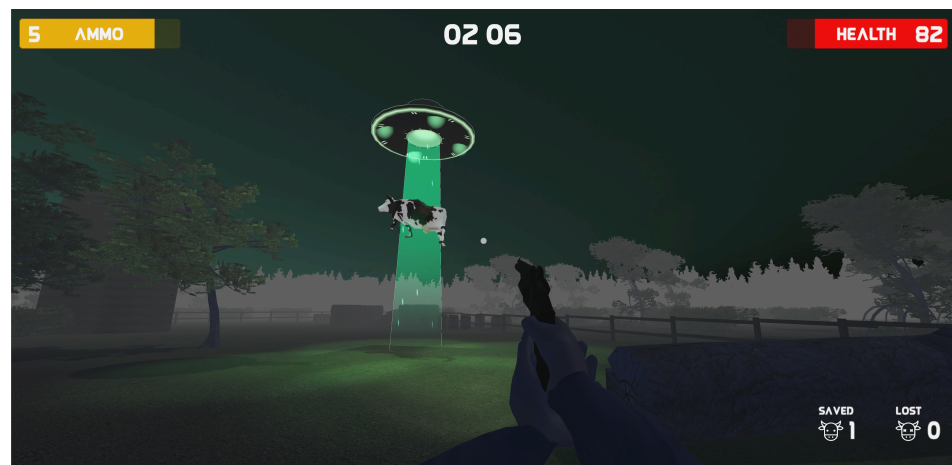
Conversely, affective-based DDA focuses on adjusting game difficulty in response to a player's emotional state, utilizing real-time physiological feedback from sensors. This method has been explored using various tools, including Electroencephalography (EEG) devices for detecting anxiety or stress [18,19] and the affective metric of focus in two-dimensional (2D) platform games [20]. Stein et al. [21] implemented affective-based DDA in a multiplayer third-person shooter, targeting excitement as the key emotional driver. Other devices, such as functional near-infrared spectroscopy (fNIRS) to assess task demand [22] and heart rate variability (HRV) for stress detection, linking higher heart rates with increased stress levels [3–5,23] to calculate the affective state of the user, have been used as well. The galvanic skin response (GSR) is another popular measure for understanding player affect in DDA research [24–26]. By gauging emotional states like focus, stress, excitement, and boredom, affective-based DDA aims to create a more immersive and emotionally resonant gaming experience.

The effectiveness of DDA, particularly when grounded in Flow Theory, has been debated. Zhixing Guo et al. [27] critiqued the reliance on Flow Theory, suggesting it leads to mixed outcomes in enhancing player experiences. Similarly, Koskinen's study indicated that situational interest increased only with significant downward difficulty adaptations [28]. These critiques underscore the complexity of applying DDA in gaming.

Recent studies have begun exploring hybrid approaches that combine performance and affective data to refine DDA [24,29]. This integration aims to address the limitations of relying solely on one type of data, marking a promising direction for future research and application in the field. Our study aims to extend this evolving discourse by comparing the efficacy of performance and emotion-based DDA with static difficulty techniques while also introducing a novel hybrid approach designed to optimize player enjoyment and engagement.

3. Game Environment: The Cattle Catcher Game

To test our Dynamic Difficulty Adjustment methods, we developed a custom game with several difficulty levels. The game, a first-person shooter titled “Cattle Catchers from Outer Space” (see Figure 2 for screenshots of the game), belongs to the endless survival genre. In this game, players aim to survive for as long as possible while defending against alien invaders. Apart from survival, players can also strive for high scores by accomplishing secondary objectives, which include escorting cows to safety, collecting crystals from defeated enemies, and protecting cows from unidentified flying objects (UFOs) aiming to abduct them. We have incorporated seven difficulty levels, ranging from extremely easy (1) to extremely hard (7), altering in-game variables associated with the enemies and the player character.



(a)



(b)

Figure 2. Screenshots of the Cattle Catchers game. (a) A UFO abducting a cow in the game. (b) Player shooting the enemies and dodging the enemy projectiles.

Design of Difficulty Levels

Before detailing the specific adjustments to player and enemy attributes across the various difficulty levels, it is important to outline the methodology behind these determinations. The values for each difficulty level were established through a series of pilot studies involving players of varying skill levels. These studies aimed to identify settings that offer a balanced and engaging experience for a wide range of players. By analyzing gameplay data and gathering feedback, we adjusted the in-game variables to ensure each difficulty level was challenging yet playable, accommodating the diverse abilities and preferences of our player base.

For the player character, three attributes vary according to difficulty: reload time, aim-down sight (ADS) speed, and aim assist. The attributes and their settings across difficulty levels (from extremely easy to extremely hard) are provided in Table 1.

- **Reload Time:** Refers to the duration required for the player character (in seconds) to reload their weapon. This duration increases with rising difficulty levels.
- **ADS Speed:** This value acts as a multiplier to the base speed for aiming down the sights, influencing how quickly the player character can enter aiming mode for the more precise targeting of enemies. A higher multiplier means quicker aiming. As the difficulty level increases, the multiplier decreases, making aiming slower and adding to the challenge.
- **Aim Assist:** This attribute determines the effective hitbox size, in meters, around each projectile fired at enemies, essentially acting as the size of the detection sphere for each shot. As the game's difficulty level increases, this detection sphere—or the aim assist—becomes smaller. This reduction requires players to be more precise with their aiming, as the margin for error decreases with higher difficulty settings.

Table 1. Difficulty adjustments for player character. Note: the reload time is measured in seconds, and the ADS speed and aim assist are measured in meters.

Difficulty	Reload Time	ADS Speed	Aim Assist
Extremely Easy	1.25	6.00	0.200
Very Easy	1.50	5.50	0.150
Easy	1.75	5.25	0.100
Medium	2.00	5.00	0.050
Hard	2.00	4.50	0.025
Very Hard	2.00	4.00	0.015
Extremely Hard	2.00	4.00	0.010

Aliens, when spawned, attempt to defeat the player. They periodically move to a location on the game map near the player and fire projectiles at the player if they have a clear line of sight and are within firing range. These enemies always target the player's center mass. However, to simplify gameplay at lower difficulties, we introduced a degree of randomness in the projectiles' trajectory. The parameters altered for these enemies and their settings across difficulty levels are detailed in Table 2.

- **Health:** This represents the amount of damage the enemy can sustain before defeat. It is specific to body shots since headshots guarantee instant kills.
- **Fire Rate:** Represents the amount of time (in seconds) that passes before the enemy can fire a new projectile.
- **Aiming Error:** Defines the inaccuracy in enemy targeting as a sphere of variable size around the player (in meters), within which enemies randomly select points to aim and fire. This mechanism introduces intentional inaccuracy in enemy attacks. The larger the sphere, the greater the inaccuracy.
- **Move Delay:** The amount of time the enemy remains stationary before repositioning closer to the player.

- **Max Spawn:** Designates the maximum number of enemies allowed on the game map simultaneously.
- **Spawn Rate:** Specifies the amount of time that passes in seconds before deploying more enemies onto the map.
- **Spawn Amount:** Indicates the number of enemies that can spawn during this timestep.

Table 2. Difficulty adjustments for alien enemies. Note: the fire rate, move delay, and spawn rate are measured in seconds; the aiming error is measured in meters.

Difficulty	Health	Fire Rate	Aiming Error	Move Delay	Max Spawn	Spawn Rate	Spawn Amount
Extremely Easy	1	1.75	0.45	5	3	10	1
Very Easy	1	1.45	0.45	4	3	10	1
Easy	2	1.20	0.45	4	6	9	1
Medium	3	1.00	0.25	2	9	8	1
Hard	3	0.85	0.20	1	12	8	2
Very Hard	3	0.75	0.15	0	15	8	2
Extremely Hard	4	0.70	0.10	0	18	8	3

UFOs, another enemy type in the game, spawn to abduct cows. Their modifiable attributes and settings across difficulty levels are presented in Table 3.

- **Health:** The amount of damage the enemy can sustain before being defeated.
- **Abduction Speed:** Acts as a multiplier to the base abduction speed of UFOs using their gravity ray. A lower value signifies a longer duration required to abduct cows.
- **Max Spawn:** Dictates the maximum UFOs present on the game map concurrently.
- **Spawn Rate:** The duration (in seconds) the game waits before spawning additional UFOs.

Table 3. Difficulty adjustments for UFO enemies. Note: the spawn rate is in seconds.

Difficulty	Health	Abd. Speed	Max Spawn	Spawn Rate
Extremely Easy	2	0.25	2	10
Very Easy	2	0.25	2	10
Easy	3	0.50	4	9
Medium	3	0.50	6	8
Hard	4	0.75	8	8
Very Hard	4	0.75	10	8
Extremely hard	6	1.00	12	8

Having outlined the game environment and the design of the difficulty levels, it is important to understand the dynamic methods by which these difficulty levels can be adjusted. In the following section, we detail the various Dynamic Difficulty Adjustment methods that were explored and tested in our research.

4. Difficulty Adjustment Methods

We implemented five difficulty adjustment methods. Two of these methods served as our baseline and employed static, unchanging difficulty levels. The other three methods leveraged either performance data, affective data (garnered from EEG and heart rate sensors), or a combination of both. To ensure more stable adjustments, both performance and affective data utilized a moving average over the last 40 s. This approach aimed to avoid abrupt difficulty fluctuations within brief periods. Each method is described in the subsequent subsections.

4.1. Fixed Difficulty

This method employs a static, medium difficulty level (4) throughout the game, serving as a control for comparative analysis against other adjustment techniques.

4.2. User Selected

Players personally select their preferred difficulty level, which can range from extremely easy (1) to extremely hard (7). While this choice often reflects a player's gaming experience, it might also be influenced by their current emotional state, such as stress or mental fatigue.

4.3. Performance-Based

Performance metrics are derived from a weighted average of diverse in-game parameters. These parameters, and their significance in gauging player skill, are listed below. The exact weights can be referenced in Table 4.

- Accuracy: This represents the player's current accuracy, measured as a percentage. A higher accuracy indicates a more skilled player.
- Headshot Ratio: This is the ratio of headshots to body shots against the alien enemy type, measured as a percentage. A higher headshot ratio suggests a more skilled player.
- Player Health: This signifies the player's current health level. This value ranges from 0 to 100.
- Last Hit: This is the time in seconds that has passed since the player last took health damage. A higher value suggests a more skilled player.
- Cows Corralled: This represents the secondary objective of successfully escorting cows to the safe zone. A higher value suggests a more skilled player.
- UFOs Destroyed: This signifies the secondary objective of destroying UFOs before they can abduct cows. A higher value suggests a more skilled player.
- Crystals Gathered: This counts the number of crystals collected, which are items dropped upon the death of the alien enemy type. A higher value suggests a more skilled player.
- Survival Time: This is the time in seconds that the player has managed to survive in the current playthrough. Lasting longer suggests a more skilled player.

Table 4. Performance-based metrics and weight values. Note: last hit and survival time are in seconds.

Points	Accuracy	Headshot Ratio	Player Health	Last Hit	Cows Corralled	UFOs Destroyed	Crystals Gathered	Survival Time
1	[0.00, 0.20)	[0.00, 0.25)	[0, 15)	[0, 1)	[0, 3)	[0, 3)	[0, 3)	[0, 24)
2	[0.20, 0.30)	[0.25, 0.35)	[15, 30)	[1, 3)	[3, 5)	[3, 5)	[3, 5)	[24, 48)
3	[0.30, 0.40)	[0.35, 0.45)	[30, 45)	[3, 6)	[5, 7)	[5, 7)	[5, 7)	[48, 72)
4	[0.40, 0.50)	[0.45, 0.55)	[45, 60)	[6, 9)	[7, 10)	[7, 10)	[7, 10)	[72, 96)
5	[0.50, 0.60)	[0.55, 0.65)	[60, 75)	[9, 12)	[10, 13)	[10, 13)	[10, 13)	[96, 120)
6	[0.60, 0.70)	[0.65, 0.75)	[75, 90)	[12, 15)	[13, 16)	[13, 16)	[13, 16)	[120, 145)
7	[0.70, 1]	[0.75, 1]	[90, 100]	[>15]	[>16]	[>16]	[>16]	[>145]
Weight	0.30	0.35	0.10	0.05	0.05	0.05	0.05	0.05

Each performance parameter is assigned a weight reflecting its significance in assessing player proficiency and expertise, with these weights derived from pilot studies and player feedback. Game performance metrics, taking into account both player skill level and the game's current state, guide the recommended difficulty level in Table 5. A weighted moving average emphasizes recent performance, focusing on the last 40 s of playtime.

Table 5. Suggested difficulty for the next timestep for both performance and emotion dynamic difficulty techniques.

Difficulty	Performance	Emotion
Extremely Easy	[0.00, 1.85)	[0.00, 1.50)
Very Easy	[1.85, 2.70)	[1.50, 2.50)
Easy	[2.70, 3.55)	[2.50, 3.50)
Medium	[3.55, 4.45)	[3.50, 4.50)
Hard	[4.45, 5.30)	[4.50, 5.50)
Very Hard	[5.30, 6.15)	[5.50, 6.50)
Extremely Hard	[≥ 6.15]	[≥ 6.50]

4.4. Emotion-Based

Drawing from Csíkszentmihályi's Flow Theory, our primary objective is to keep players within an optimal mental state that balances stress and engagement, thereby preventing either boredom or excessive mental strain. Unlike performance-based DDA, this method does not gauge the player's skill. Instead, it harnesses physiological sensors to gauge players' stress and engagement levels directly, adjusting the game difficulty accordingly.

Our system aggregates stress and engagement metrics, each ranging from 0 to 100, into a single metric. To ensure this combined metric is compatible with the 0 to 7 scale used in our performance-based Dynamic Difficulty Adjustment (DDA) system, we normalized it by dividing the sum by 28.57. This divisor was calculated by dividing the maximum possible combined value of 200 (the sum of the maximums for both metrics) by the maximum of the target range, 7, to seamlessly integrate with other DDA calculations. The normalized value then influences the moving average formula, which determines the overall game difficulty based on the last 40 s of gameplay, providing a buffer against abrupt difficulty changes. This normalization process and its application in adjusting game difficulty are further detailed in Table 5 and Figure 3.

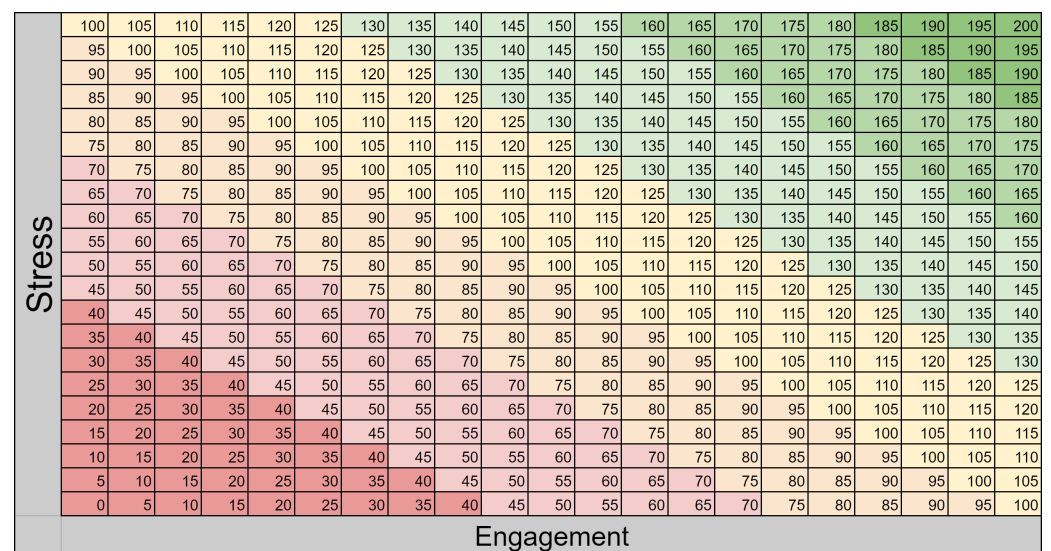


Figure 3. Chart of emotion-based difficulty adjustment. Stress and engagement scores are merged to form a unified emotion metric, which is categorized into seven distinct levels, corresponding to suggested difficulty levels (1–7) for the upcoming timestep. Each colored square represents one of the seven difficulty levels suggested for the next timestep.

4.5. Combined Approach

This Dynamic Difficulty Adjustment (DDA) framework employs a hybrid strategy, integrating both performance-based and emotion-based metrics to tailor the game's challenge in real-time. To achieve this, we computed a weighted average, where the emotion-based metric receives a weighting of 0.6 and the performance-based metric a weighting of 0.4.

This weighting scheme was carefully chosen based on insights gained from pilot studies, which emphasized the significance of the player's current emotional state in enhancing game engagement and overall experience.

In our pilot studies, we observed that players' enjoyment and immersion were more deeply impacted by adjustments that considered their emotional responses, such as stress and engagement levels, compared to purely performance-based adjustments. Therefore, by assigning a higher weight to the emotion-based metric, we aimed to ensure that the game dynamically adapts in a manner that prioritizes player sentiment, fostering a more personalized and emotionally resonant gaming experience. This approach allows for a nuanced balance, ensuring that while performance factors remain important, they do not overshadow the importance of maintaining an engaging and emotionally satisfying challenge level for the player.

5. Methodology

5.1. Goals

This study aims to explore the intricacies of player preferences regarding game difficulty settings, specifically examining the impact of DDA compared to static difficulty settings. Understanding player preferences can significantly enhance user experience and engagement. From this, we propose the following research questions:

RQ1: Do participants prefer DDA over static difficulties in video games?

RQ2: Among DDA techniques, which do participants prefer: emotion-based, performance-based, or a combination of techniques?

RQ3: Are higher difficulty levels within DDA techniques associated with increased stress and engagement among players?

RQ4: What differences are observed in performance and physiological metrics when applying various DDA techniques?

5.2. Experimental Design and Procedure

Upon entering the laboratory, participants were escorted to the experimental area. They then completed several forms: a consent form, a COVID-19 precautionary questionnaire, and a pre-experiment questionnaire concerning gaming experience and immersive tendencies. Following this, participants were fitted with physiological sensors. We ensured that the devices were correctly positioned and provided reliable readings. A training session was then held to acquaint participants with the game's controls, environment, and objectives. After training, they entered a baseline data collection phase, where they were instructed to relax with their eyes open for thirty seconds, followed by an equivalent period with their eyes closed. This allowed us to capture essential baseline parameters like resting heart rate and baseline emotional states.

Participants were outfitted with an EEG device while playing the shooter game and wore a wristwatch that measured HRV. These devices monitored the players' affective states, allowing the game's difficulty to dynamically adjust in response. Additionally, we recorded in-game performance metrics to modify the game in real-time and for post-experiment statistical analysis.

Participants then underwent testing for the five different difficulty adjustment methods, with the sequence determined using a Latin square. Each method was tested either until a three-minute timer expired or the game character died by losing all health due to enemy projectiles. Post-technique questionnaires (detailed in Table 6) were completed after each method to complement the EEG data and in-game metrics with participants' feedback. Following the testing of all five methods, participants were offered a break. A second trial, repeating all procedures, was conducted post-break.

Table 6. Post-technique questionnaire. Participants answered questions regarding their current affective states during the playthrough. Answers are in the form of a 7-point Likert scale.

Post Technique Questionnaire	
Q1	Please rate your average stress level while playing this iteration of the game
Q2	Please rate your average excitement level while playing this iteration of the game
Q3	Please rate your average focus level while playing this iteration of the game
Q4	How much fun did you have playing the game this iteration?
Q5	How difficult would you rate this iteration of the game?
Q6	How much effort did you put into playing the game?
Q7	How well do you think you performed in the game?
Q8	Did you feel that you were trying your best?
Q9	To what extent did you enjoy playing the game, rather than something you were just doing?
Q10	Was there a noticeable difference in difficulty from the previous trial/trials?

5.3. Participants

In our study, we recruited a total of 31 participants from the university population, comprising 27 males and 4 females, with ages ranging from 19 to 33 years (mean: 22.46). To categorize participants based on their gaming experience, we administered a pre-experiment questionnaire containing inquiries such as ‘Do you currently play games’, ‘How often do you play games?’, ‘How long have you been playing games?’, and a self-reported assessment of skill level in video games.

Based on their responses to these questions, participants were divided into two distinct categories: casual and experienced players. Casual players were identified as those who reported either limited gaming activity or no gaming experience at all. On the other hand, experienced players were defined as those who indicated frequent gaming activity and self-reported proficiency in video games. Of the 31 participants, 14 were classified as casual or non-gamers, while the remaining 17 were categorized as experienced gamers. This categorization facilitated subsequent analysis under the label expertise, allowing us to explore the potential differences in gaming preferences and experiences between these two groups.

5.4. Equipment and Software

The experimental setup, illustrated in Figure 4, comprised a 27" Dell monitor, an Emotiv Insight EEG device, a Polar Verity optical heart rate sensor wristwatch, and a PC with an Intel Core i7 8700K processor, GTX 1080 Graphics card, and 16 GB RAM, running on Microsoft Windows 10. The game was developed using Unity 3D v2018.4.14f1.

The Emotiv Insight, a 5-channel low-power EEG device, provided processed emotion values such as Stress, Engagement, Interest, Excitement, Focus, and Relaxation every 10 s. Definitions for these emotions can be found on the Emotiv website [30]. This study primarily utilized the Stress and Engagement metrics for the emotion-based and Combined techniques.

The Polar Verity Sense wristwatch measured heart rate variability, aiding in the determination of participant stress levels. Prior research has linked heart rate variability with stress, finding that elevated stress and anxiety levels correspond to decreased intervals between heartbeats [3–5]. For simplicity, we recorded participants’ mean heart rate for each technique for that timestep and compared it to their resting heart rate during the baseline data collection. The Polar Verity Sense provided heart rate variability data at a rate of 4 Hz.

Statistical analysis, using IBM SPSS 26, was conducted using a variety of techniques tailored to the nature of the collected data. For the game and physiological data, we performed a one-way multivariate analysis of variance (MANOVA) to examine the overall differences between techniques, independent samples T-tests to compare means between casual and experienced players, and pairwise comparisons for dependent samples T-tests

to assess changes within groups across different conditions. Additionally, we used Holm's sequential Bonferroni adjustment to correct for type I errors [31].



Figure 4. Experimental setup.

For the post-questionnaire data, which utilized a 7-point Likert scale, we employed the Friedman Test to analyze differences across multiple related groups, Mann–Whitney U Test to compare two independent groups on a non-parametric variable, and Wilcoxon Signed-Rank Test to compare paired samples on a non-parametric variable.

Additionally, visualizations of the data were created using Python using the matplotlib and pandas libraries for enhanced interpretation and presentation.

6. Results

6.1. Game Data and Physiological Sensors

In this study, we evaluated the effectiveness of various DDA techniques. Our approach involved collecting an array of data, which encompassed both sensor outputs and in-game metrics, as well as feedback from participants through post-experiment questionnaires to determine which DDA techniques were most successful.

In the context of the varying levels of difficulty, it was found that the Emotion technique posed the highest challenge, with an average difficulty rating of 5.93 as shown in Figure 5. On the other end of the spectrum, Performance presented the least degree of difficulty, averaging at 3.92. The Combined technique exhibited a middle-ground difficulty level, with its average score hovering around 4.71. When users were given the liberty to choose their level of difficulty, the self-selected average turned out to be 4.29. The difficulty scale in this study was designed to range from 1 to 7.

When analyzing the mean difficulty levels across different DDA techniques, accounting for expertise, a significant difference between casual and experienced participants is observable during the Performance technique. Specifically, the average difficulties, over all the techniques, for casual and experienced participants were 3.43 and 4.33, respectively, as depicted in Figure 6. A similar trend emerged in the User Selected technique, where casual players opted for lower difficulties on average. Here, casual players generally preferred a moderate difficulty setting, averaging 4.00, in contrast to the more experienced players who leaned towards a slightly tougher challenge, with an average difficulty of 4.53. This pattern is also evident in the Combined and Emotion techniques, where more seasoned players faced marginally higher difficulties on average overall as well.

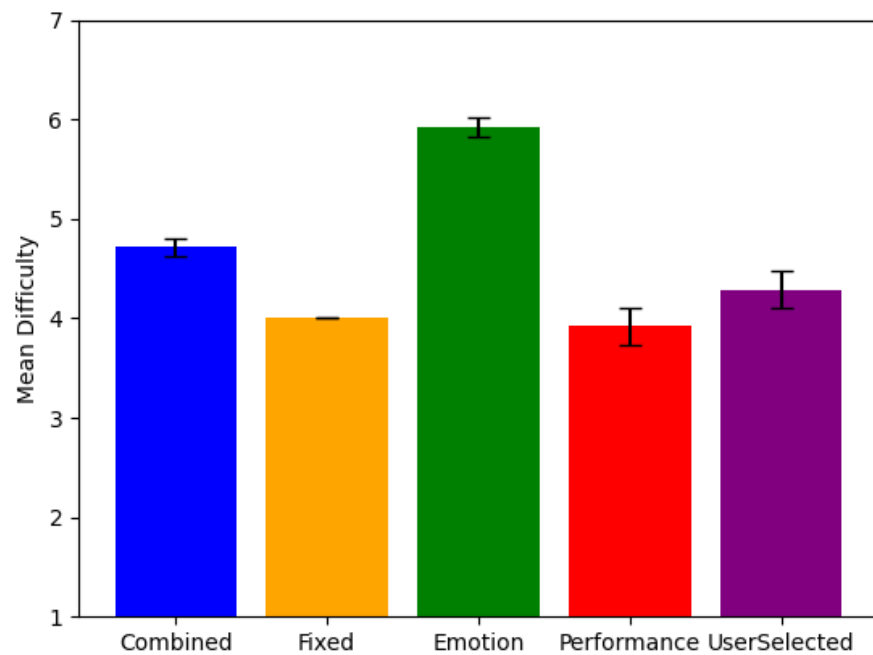


Figure 5. Mean difficulty per technique.

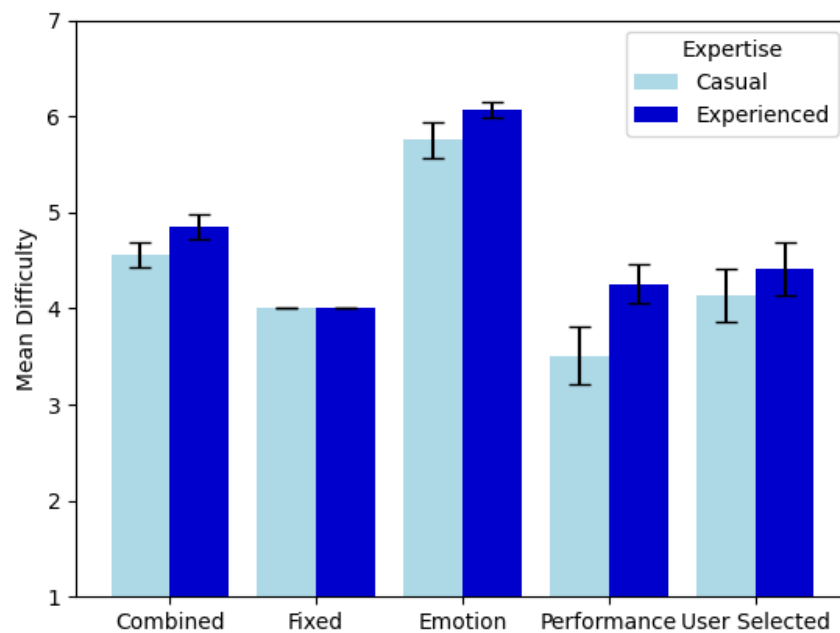


Figure 6. Mean difficulty per technique and expertise.

In Figure 7, we present the mean values for an array of game data alongside the physiological sensor measurements obtained during our user study. It is important to emphasize that, in contrast to initial expectations, the variability observed in most physiological sensor readings, the EEG and heart rate data, was significantly lower than anticipated. For example, the standard deviation across all DDA techniques indicates minimal variability in the Engagement levels, with standard deviation values ranging from 0.058 (Performance) to 0.072 (Combined). This uniformity contrasts with our initial expectations of a wider variability in emotional responses. This trend of limited variance is consistent across all emotions captured by the EEG. In terms of HRV, the Emotion technique was unique by eliciting a notable increase in heart rate. This was determined by calculating the heart rate

(HR) difference and comparing the average heart rate at each timestep for each technique to the baseline resting heart rate established during the initial data collection phase.

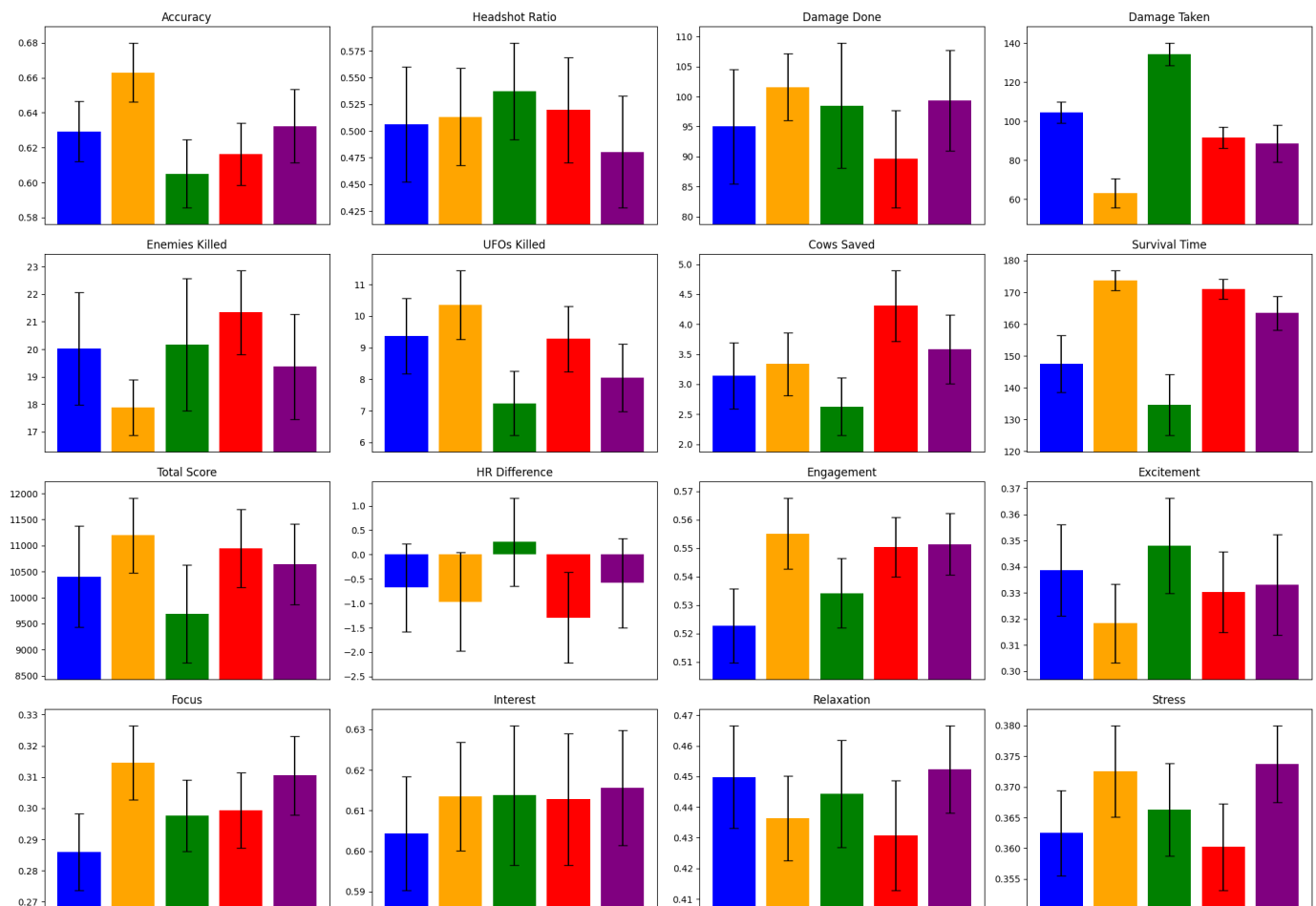


Figure 7. Mean values of game data and physiological sensor readings across different techniques. Techniques are color-coded as follows: Combined (blue), Fixed (yellow), Emotion (green), Performance (red), and User Selected (purple).

In evaluating performance metrics, the impact of the Emotion technique on gameplay becomes evident through its consistent under-performance in several key areas: it notably resulted in fewer UFOs killed, fewer cows saved, shorter survival times, and lower total scores, alongside recording the highest average damage taken. In contrast, the Fixed technique demonstrated higher overall readings, with higher total scores, more UFOs killed, longer survival times, greater damage done, less damage taken, and higher accuracy.

To provide a more comprehensive understanding of these findings, a detailed statistical analysis will be conducted on these recorded performance metrics and physiological data. For our sensor and in-game data, we performed a series of statistical tests. Our analysis includes a MANOVA for evaluating game metrics impacts (Table 7), and T-tests to explore differences between groups and techniques, detailed in Table 8 for independent samples and Tables 9 and 10 for dependent samples. Through this analysis, we aim to illustrate the intricate ways DDA shapes player experience and game dynamics.

Table 7. MANOVA revealing effects of technique type and gaming expertise on participant performance.

Metric	Technique	Technique × Expertise
Difficulty	$F_{4,26} = 108.430, p < 0.050$	$F_{4,26} = 3.483, p < 0.050$
Accuracy	$F_{4,26} = 3.337, p < 0.050$	$F_{4,26} = 0.311, p = 0.868$
Headshot Ratio	$F_{4,26} = 0.485, p = 0.746$	$F_{4,26} = 0.038, p = 0.997$
Damage Done	$F_{4,26} = 4.013, p < 0.050$	$F_{4,26} = 1.549, p = 0.217$
Damage Taken	$F_{4,26} = 18.761, p < 0.050$	$F_{4,26} = 2.813, p < 0.050$
Enemies Killed	$F_{4,26} = 3.856, p < 0.050$	$F_{4,26} = 1.635, p = 0.196$
UFOs Killed	$F_{4,26} = 5.403, p < 0.050$	$F_{4,26} = 0.484, p = 0.747$
Survival Time	$F_{4,26} = 7.782, p < 0.050$	$F_{4,26} = 2.021, p = 0.121$
Cows Saved	$F_{4,26} = 5.479, p < 0.050$	$F_{4,26} = 0.675, p = 0.615$
Total Score	$F_{4,26} = 3.440, p < 0.050$	$F_{4,26} = 0.560, p = 0.693$
HR Difference	$F_{4,26} = 1.633, p = 0.196$	$F_{4,26} = 0.652, p = 0.631$
Engagement	$F_{4,26} = 2.091, p = 0.111$	$F_{4,26} = 1.557, p = 0.215$
Excitement	$F_{4,26} = 0.873, p = 0.459$	$F_{4,26} = 0.383, p = 0.819$
Focus	$F_{4,26} = 1.508, p = 0.229$	$F_{4,26} = 5.719, p < 0.050$
Interest	$F_{4,26} = 0.946, p = 0.453$	$F_{4,26} = 2.886, p < 0.050$
Relaxation	$F_{4,26} = 0.778, p = 0.550$	$F_{4,26} = 3.528, p < 0.050$
Stress	$F_{4,26} = 1.193, p = 0.337$	$F_{4,26} = 0.914, p < 0.050$

Our one-way multivariate analysis of variance investigated the impacts of different difficulty adjustment techniques—emotion-based, performance-based, a combined approach, fixed difficulty, and user selected—on game performance metrics and physiological sensor data. The findings reveal significant differences in how these adjustment techniques influence key performance metrics, such as difficulty, accuracy, and total score, showcasing the distinct impact each method has on the gaming experience (for detailed statistics, see Table 7).

When analyzing the data by separating players based on expertise (casual vs. experienced), the overall impact of expertise was less pronounced, with only difficulty and damage taken showing significant differences. This suggests that while the type of difficulty adjustment technique plays an important role across all players, the division of expertise level does not uniformly affect most game performance metrics. However, physiological measures such as focus, interest, relaxation, and stress revealed significant differences when players were categorized by expertise, highlighting how expertise influences players' emotional and cognitive reactions to gameplay.

Further exploring the differences between casual and experienced players, independent samples T-tests were conducted to assess the effectiveness of five DDA techniques in modifying key game performance metrics and physiological responses, considering player expertise. Due to multiple conditions tested simultaneously, we used Holm's sequential Bonferroni adjustment to correct for type I errors [31]. Significant differences between casual and experienced players were observed in the following areas as detailed in Table 8:

- **Accuracy:** Significant differences in accuracy were found between casual and experienced players for the Combined and Performance DDA techniques. Specifically, for the Combined technique, casual players had a mean accuracy of 0.589, while experienced players had a mean of 0.663. Similarly, under the Performance technique, the mean accuracy was 0.575 for casual players and 0.650 for experienced players.
- **Damage Done:** For the Emotion DDA technique, significant differences were observed between casual and experienced players. Casual players had a mean damage done of 75.03, whereas experienced players had a higher mean of 117.85, indicating that experienced players tend to inflict more damage within this technique.
- **Damage Taken:** Only the Fixed technique showed significant differences when comparing casual and experienced players. Casual players had a mean damage taken of 77.25, while experienced players had a significantly lower mean of 51.35, indicating

that experienced players manage to avoid taking damage more effectively than casual players under this static difficulty.

- **Survival Time:** For the Emotion DDA technique, significant differences were observed between casual and experienced players. Casual players had a mean survival time of 115 s, while experienced players had a mean of 150 s, indicating that experienced players tended to survive longer under this technique.
- **Total Score:** Significant variation was found for the Performance technique when comparing casual and experienced players. Casual players had a mean total score of 9180, while experienced players had a significantly higher mean of 12,397, indicating that experienced players generally achieve higher scores under this condition.
- **Physiological Sensor Data:** Minimal significant differences were found when accounting for expertise, except for the Focus metric with the User Selected technique, where a significant difference was noted. Casual players had a mean focus level of 0.282, while experienced players had a higher mean of 0.334, indicating that experienced players tend to maintain better focus under this condition.

Table 8. Comparison of game performance metrics and physiological responses between casual and experienced players across five difficulty adjustment techniques using independent samples T-tests.

Metric	Combined (CB)	Fixed (FX)	Emotion (EM)	Performance (PE)	User Selected (US)
Difficulty	$t_{29} = -1.738$, $p = 0.093$	N/A	$t_{29} = -1.187$, $p = 0.245$	$t_{29} = -2.710$, $p = 0.011$	$t_{29} = -1.391$, $p = 0.175$
Accuracy	$t_{29} = -3.338$, $p < 0.010$	$t_{29} = -2.446$, $p = 0.021$	$t_{29} = -1.772$, $p = 0.087$	$t_{29} = -2.880$, $p < 0.010$	$t_{29} = -1.850$, $p = 0.075$
Headshot Ratio	$t_{29} = -1.904$, $p = 0.067$	$t_{29} = -2.012$, $p = 0.054$	$t_{29} = -2.022$, $p = 0.052$	$t_{29} = -1.956$, $p = 0.060$	$t_{29} = -1.884$, $p = 0.070$
Damage Done	$t_{29} = -2.509$, $p = 0.018$	$t_{29} = -2.326$, $p = 0.027$	$t_{29} = -2.859$, $p < 0.010$	$t_{29} = -2.715$, $p = 0.011$	$t_{29} = -2.345$, $p = 0.026$
Damage Taken	$t_{29} = 1.442$, $p = 0.160$	$t_{29} = 3.564$, $p < 0.010$	$t_{29} = -1.136$, $p = 0.265$	$t_{29} = 1.276$, $p = 0.212$	$t_{29} = 0.768$, $p = 0.449$
Enemies Killed	$t_{29} = -2.708$, $p = 0.011$	$t_{29} = -1.499$, $p = 0.145$	$t_{29} = -2.338$, $p = 0.026$	$t_{29} = -1.780$, $p = 0.086$	$t_{29} = -1.778$, $p = 0.086$
UFOs Killed	$t_{29} = -1.448$, $p = 0.158$	$t_{29} = -1.864$, $p = 0.072$	$t_{29} = -1.555$, $p = 0.131$	$t_{29} = -2.158$, $p = 0.039$	$t_{29} = -1.366$, $p = 0.183$
Cows Saved	$t_{29} = -1.775$, $p = 0.086$	$t_{29} = -2.667$, $p = 0.012$	$t_{29} = -2.108$, $p = 0.044$	$t_{29} = -1.677$, $p = 0.104$	$t_{29} = -1.273$, $p = 0.213$
Survival Time	$t_{29} = -2.615$, $p = 0.014$	$t_{29} = -2.385$, $p = 0.024$	$t_{29} = -3.033$, $p < 0.010$	$t_{29} = -0.877$, $p = 0.388$	$t_{29} = -1.034$, $p = 0.310$
Total Score	$t_{29} = -2.617$, $p = 0.039$	$t_{29} = -2.410$, $p = 0.023$	$t_{29} = -2.635$, $p = 0.013$	$t_{29} = -2.777$, $p < 0.010$	$t_{29} = -2.038$, $p = 0.051$
HR Difference	$t_{29} = 0.037$, $p = 0.971$	$t_{29} = -0.066$, $p = 0.948$	$t_{29} = 0.292$, $p = 0.772$	$t_{29} = -0.383$, $p = 0.704$	$t_{29} = -0.148$, $p = 0.883$
Engagement	$t_{29} = -0.622$, $p = 0.539$	$t_{29} = -0.072$, $p = 0.943$	$t_{29} = 1.114$, $p = 0.275$	$t_{29} = 1.528$, $p = 0.137$	$t_{29} = -1.341$, $p = 0.190$
Excitement	$t_{29} = 0.842$, $p = 0.407$	$t_{29} = 1.621$, $p = 0.116$	$t_{29} = 0.413$, $p = 0.683$	$t_{29} = 1.400$, $p = 0.172$	$t_{29} = 1.047$, $p = 0.304$
Focus	$t_{29} = -0.830$, $p = 0.413$	$t_{29} = -1.706$, $p = 0.099$	$t_{29} = 1.696$, $p = 0.101$	$t_{29} = 1.020$, $p = 0.316$	$t_{29} = -3.011$, $p < 0.010$
Interest	$t_{29} = -0.501$, $p = 0.620$	$t_{29} = 0.800$, $p = 0.430$	$t_{29} = -1.022$, $p = 0.315$	$t_{29} = -0.724$, $p = 0.475$	$t_{29} = 0.583$, $p = 0.565$
Relaxation	$t_{29} = -0.194$, $p = 0.848$	$t_{29} = 1.450$, $p = 0.158$	$t_{29} = -1.467$, $p = 0.153$	$t_{29} = -1.324$, $p = 0.196$	$t_{29} = 1.933$, $p = 0.063$
Stress	$t_{29} = -1.001$, $p = 0.325$	$t_{29} = -0.460$, $p = 0.649$	$t_{29} = 0.110$, $p = 0.913$	$t_{29} = -0.762$, $p = 0.452$	$t_{29} = -1.394$, $p = 0.174$

Table 9. Pair-wise comparisons of dependent samples T-test across metrics (pairs 1–5).

Metric/Pair	CB × FX	CB × EM	CB × PE	CB × US	PE × US
Difficulty	$t_{30} = 7.553$, $p < 0.005$	$t_{30} = -8.935$, $p < 0.005$	$t_{30} = 6.042$, $p < 0.005$	$t_{30} = 2.193$, $p = 0.036$	$t_{30} = -1.551$, $p = 0.131$
Accuracy	$t_{30} = -2.476$, $p = 0.019$	$t_{30} = 1.427$, $p = 0.164$	$t_{30} = 0.833$, $p = 0.412$	$t_{30} = -0.177$, $p = 0.860$	$t_{30} = -0.761$, $p = 0.453$
Headshot Ratio	$t_{30} = -0.231$, $p = 0.819$	$t_{30} = -0.909$, $p = 0.371$	$t_{30} = -0.422$, $p = 0.676$	$t_{30} = 0.839$, $p = 0.408$	$t_{30} = 0.943$, $p = 0.353$
Damage Done	$t_{30} = -1.239$, $p = 0.225$	$t_{30} = -0.704$, $p = 0.487$	$t_{30} = 1.105$, $p = 0.278$	$t_{30} = -0.678$, $p = 0.503$	$t_{30} = -1.516$, $p = 0.140$
Damage Taken	$t_{30} = 4.695$, $p < 0.005$	$t_{30} = -4.597$, $p < 0.005$	$t_{30} = 2.460$, $p = 0.020$	$t_{30} = 2.042$, $p = 0.050$	$t_{30} = 0.296$, $p = 0.769$
Enemies Killed	$t_{30} = 1.454$, $p = 0.156$	$t_{30} = -0.094$, $p = 0.926$	$t_{30} = -0.886$, $p = 0.383$	$t_{30} = 0.416$, $p = 0.680$	$t_{30} = 1.186$, $p = 0.245$
UFOs Killed	$t_{30} = -1.677$, $p = 0.104$	$t_{30} = 3.176$, $p < 0.005$	$t_{30} = 0.210$, $p = 0.835$	$t_{30} = 1.734$, $p = 0.093$	$t_{30} = 1.800$, $p = 0.082$
Cows Saved	$t_{30} = -0.623$, $p = 0.538$	$t_{30} = 2.147$, $p = 0.040$	$t_{30} = -3.321$, $p < 0.005$	$t_{30} = -1.401$, $p = 0.172$	$t_{30} = 1.811$, $p = 0.080$
Survival Time	$t_{30} = -3.598$, $p < 0.005$	$t_{30} = 2.803$, $p = 0.009$	$t_{30} = -2.753$, $p = 0.010$	$t_{30} = -2.139$, $p = 0.041$	$t_{30} = 1.420$, $p = 0.166$
Total Score	$t_{30} = -1.584$, $p = 0.124$	$t_{30} = 1.689$, $p = 0.102$	$t_{30} = -1.245$, $p = 0.223$	$t_{30} = -0.454$, $p = 0.653$	$t_{30} = 0.824$, $p = 0.416$
HR Difference	$t_{30} = -0.401$, $p = 0.692$	$t_{30} = 2.139$, $p = 0.041$	$t_{30} = -1.353$, $p = 0.186$	$t_{30} = 0.334$, $p = 0.741$	$t_{30} = 1.496$, $p = 0.145$
Engagement	$t_{30} = -2.316$, $p = 0.028$	$t_{30} = -0.734$, $p = 0.469$	$t_{30} = -1.952$, $p = 0.060$	$t_{30} = -2.163$, $p = 0.039$	$t_{30} = -0.056$, $p = 0.956$
Excitement	$t_{30} = 1.226$, $p = 0.230$	$t_{30} = -0.725$, $p = 0.474$	$t_{30} = 0.724$, $p = 0.475$	$t_{30} = 0.444$, $p = 0.660$	$t_{30} = -0.204$, $p = 0.840$
Focus	$t_{30} = -2.333$, $p = 0.027$	$t_{30} = -0.890$, $p = 0.381$	$t_{30} = -0.995$, $p = 0.328$	$t_{30} = -1.817$, $p = 0.079$	$t_{30} = -0.670$, $p = 0.508$
Interest	$t_{30} = -1.052$, $p = 0.301$	$t_{30} = -0.737$, $p = 0.467$	$t_{30} = -0.673$, $p = 0.506$	$t_{30} = -1.411$, $p = 0.169$	$t_{30} = -0.231$, $p = 0.819$
Stress	$t_{30} = -1.605$, $p = 0.119$	$t_{30} = -0.647$, $p = 0.523$	$t_{30} = 0.494$, $p = 0.625$	$t_{30} = -1.828$, $p = 0.078$	$t_{30} = -2.350$, $p = 0.026$
Relaxation	$t_{30} = 0.911$, $p = 0.370$	$t_{30} = 0.338$, $p = 0.738$	$t_{30} = 1.184$, $p = 0.246$	$t_{30} = -0.148$, $p = 0.883$	$t_{30} = -1.191$, $p = 0.243$

We conducted a dependent samples T-test analysis for pairwise comparisons to measure the impact of various DDA techniques on different metrics as shown in Tables 9 and 10. These comparisons allowed us to discern the differential effects between pairs of techniques on the same metric, shedding light on how each specific DDA approach influences these aspects. For the purposes of this analysis, we utilized the following abbreviations to denote specific techniques: CB for Combined, FX for Fixed, EM for Emotion, PE for Performance, and US for User Selected. Below, we summarize the combinations of techniques that demonstrated significant difference for each metric. Since we tested multiple pairs, we used Holm’s sequential Bonferroni adjustment to correct for type I errors [31].

- Difficulty: Emotion exhibited higher difficulty overall compared to other techniques. The mean difficulty score were 5.916 for Emotion, significantly higher than Combined (4.704), Fixed (4.000), Performance (3.895), and User Selected (4.282). The Combined technique (4.704) was more difficult than Fixed (4.00) and Performance (3.895).
- Accuracy: The Fixed technique showed a higher mean accuracy (0.662) compared to Emotion (0.603).
- Damage Taken: The Emotion technique resulted in a higher mean damage taken (133.890) compared to Combined (104.862), Fixed (63.884), Performance (91.707), and User Selected (89.209). Significant differences were noted across all pairwise comparisons involving Emotion. In addition, the Fixed (63.884) technique showed lower

levels of damage taken when compared to Combined (104.371) and Performance (91.452).

- Enemies Killed: Performance led to a higher mean enemies killed (21.176) compared to Fixed (17.761), demonstrating a statistically significant difference.
- UFOs Killed: Combined (9.371), Fixed (10.277), and Performance (9.178) techniques showed higher means in UFOs destroyed compared to Emotion (7.180), with significant differences noted in pairwise comparisons. The Fixed (10.355) technique also exhibited higher amounts of UFOs destroyed compared to User Selected (8.048).
- Cows Saved: The Performance technique showed a higher mean (4.250) in cows saved compared to Combined (3.092), Fixed (3.271), and Emotion (2.569).
- Survival Time: Emotion had a lower mean survival time (133.471) compared to other techniques, with significant decreases noted against Fixed (173.308), Performance (170.854), and User Selected (163.095). Additionally, the Combined technique (146.511) showed a lower survival time compared to Fixed (173.308).
- Total Score: Only the FX×EM ($t_{30} = 3.184$, $p < 0.005$) and EM×PE ($t_{30} = -3.300$, $p < 0.005$) pairs presented substantial differences in total scores. Both Fixed and Performance techniques exhibited higher total scores compared to Emotion.
- Total Score: The Fixed (10,643) and Performance (11,340) techniques yielded higher total scores compared to Emotion (9478) in pairwise comparisons.

Table 10. Pair-wise comparisons of dependent samples T-test across metrics (pairs 6–10).

Metric/Pair	FX × EM	FX × PE	FX × US	EM × PE	EM × US
Difficulty	$t_{30} = -19.974$, $p < 0.005$	$t_{30} = 0.445$, $p = 0.659$	$t_{30} = -1.510$, $p = 0.142$	$t_{30} = 8.926$, $p < 0.005$	$t_{30} = 7.649$, $p < 0.005$
Accuracy	$t_{30} = 3.315$, $p < 0.005$	$t_{30} = 2.892$, $p = 0.007$	$t_{30} = 1.690$, $p = 0.101$	$t_{30} = -0.567$, $p = 0.575$	$t_{30} = -1.411$, $p = 0.168$
Headshot Ratio	$t_{30} = -1.008$, $p = 0.322$	$t_{30} = -0.280$, $p = 0.781$	$t_{30} = 0.953$, $p = 0.348$	$t_{30} = 0.700$, $p = 0.489$	$t_{30} = 1.416$, $p = 0.167$
Damage Done	$t_{30} = 0.440$, $p = 0.663$	$t_{30} = 2.662$, $p = 0.012$	$t_{30} = 0.451$, $p = 0.655$	$t_{30} = 2.032$, $p = 0.051$	$t_{30} = -0.102$, $p = 0.919$
Damage Taken	$t_{30} = -7.049$, $p < 0.005$	$t_{30} = -3.106$, $p < 0.005$	$t_{30} = -2.360$, $p = 0.025$	$t_{30} = 6.287$, $p < 0.005$	$t_{30} = 4.948$, $p < 0.005$
Enemies Killed	$t_{30} = -1.207$, $p = 0.237$	$t_{30} = -3.285$, $p < 0.005$	$t_{30} = -0.990$, $p = 0.330$	$t_{30} = -0.891$, $p = 0.380$	$t_{30} = 0.379$, $p = 0.707$
UFOs Killed	$t_{30} = 4.987$, $p < 0.005$	$t_{30} = 2.453$, $p = 0.020$	$t_{30} = 3.140$, $p < 0.005$	$t_{30} = -3.721$, $p < 0.005$	$t_{30} = -1.201$, $p = 0.239$
Cows Saved	$t_{30} = 2.484$, $p = 0.019$	$t_{30} = -3.620$, $p < 0.005$	$t_{30} = -0.615$, $p = 0.543$	$t_{30} = -4.714$, $p < 0.005$	$t_{30} = -2.816$, $p = 0.009$
Survival Time	$t_{30} = 5.077$, $p < 0.005$	$t_{30} = 0.671$, $p = 0.508$	$t_{30} = 1.967$, $p = 0.058$	$t_{30} = -4.105$, $p < 0.005$	$t_{30} = -3.500$, $p < 0.005$
Total Score	$t_{30} = 3.184$, $p < 0.005$	$t_{30} = 0.794$, $p = 0.434$	$t_{30} = 1.800$, $p = 0.082$	$t_{30} = -3.300$, $p < 0.005$	$t_{30} = -1.851$, $p = 0.074$
HR Difference	$t_{30} = 1.886$, $p = 0.069$	$t_{30} = -0.661$, $p = 0.514$	$t_{30} = 0.634$, $p = 0.531$	$t_{30} = -2.728$, $p = 0.011$	$t_{30} = -2.040$, $p = 0.050$
Engagement	$t_{30} = 1.127$, $p = 0.269$	$t_{30} = 0.390$, $p = 0.699$	$t_{30} = 0.255$, $p = 0.800$	$t_{30} = -1.242$, $p = 0.224$	$t_{30} = -1.360$, $p = 0.184$
Excitement	$t_{30} = -1.965$, $p = 0.059$	$t_{30} = -1.001$, $p = 0.325$	$t_{30} = -1.026$, $p = 0.313$	$t_{30} = 1.598$, $p = 0.120$	$t_{30} = 0.907$, $p = 0.372$
Focus	$t_{30} = 0.991$, $p = 0.330$	$t_{30} = 1.317$, $p = 0.198$	$t_{30} = 0.290$, $p = 0.774$	$t_{30} = -0.129$, $p = 0.898$	$t_{30} = -0.904$, $p = 0.373$
Interest	$t_{30} = -0.021$, $p = 0.984$	$t_{30} = 0.060$, $p = 0.952$	$t_{30} = -0.221$, $p = 0.827$	$t_{30} = 0.095$, $p = 0.925$	$t_{30} = -0.132$, $p = 0.896$
Stress	$t_{30} = 1.061$, $p = 0.297$	$t_{30} = 1.855$, $p = 0.073$	$t_{30} = -0.242$, $p = 0.810$	$t_{30} = 1.092$, $p = 0.283$	$t_{30} = -1.307$, $p = 0.201$
Relaxation	$t_{30} = -0.428$, $p = 0.672$	$t_{30} = 0.349$, $p = 0.729$	$t_{30} = -1.206$, $p = 0.237$	$t_{30} = 0.992$, $p = 0.329$	$t_{30} = -0.464$, $p = 0.646$

6.2. Post-Questionnaire Results

Understanding and integrating user feedback directly is important when designing systems to ensure the readings collected from the physiological sensors match the expected outcomes. After testing each technique, participants were requested to complete a questionnaire. The purpose of this questionnaire was to collect self-reported data about the participants' emotional states and performance. This additional information was intended to complement the data gathered through physiological sensors, thereby providing a more comprehensive understanding of the participants' emotional states during the trial. Table 6 presents the questionnaire in full and the specific questions asked. Responses were collected on a 7-point Likert Scale.

Analysis of participant responses from Figure 8 offers insights into their experiences across various Dynamic Difficulty Adjustment (DDA) techniques. Here is a brief overview based on the specific questions asked:

- Stress (Q1): The Emotion technique was reported as the most stressful, with an average score of 4.11, whereas the Fixed technique was seen as least stressful, scoring 3.42 on average.
- Excitement (Q2): Self-reported excitement scores were relatively close across techniques, with the Emotion technique slightly leading at 4.98, and the Fixed technique at the lower end with 4.63.
- Focus (Q3): The Performance technique with an average of 5.63 was highlighted for requiring the highest focus, although differences among techniques were minimal.
- Fun (Q4): Self-reported enjoyment favored dynamic difficulties (Combined at 5.11, Emotion at 5.11, Performance at 5.15) over static ones (Fixed at 4.89, User Selected at 5.02).
- Perceived Difficulty (Q5): Reflecting the actual game difficulty, participants found the Emotion technique to be the hardest, with a mean score of 5.34, followed in descending order by Combined (4.60), User Selected (4.45), Performance (4.08), and Fixed (3.84).
- Effort (Q6): Participants reported putting more effort into sessions with dynamic difficulties, particularly with Combined (5.65), Emotion (5.63), and Performance (5.47) techniques.
- Performance (Q7): When asked to self-rate performance participants felt they performed best on the Performance technique with an average response value of 5.19 and their response for worst performance being the Emotion technique with an average response value of 3.81.
- Trying Best (Q8): Similar to effort, players reported trying harder when presented with the a dynamic technique over a static one.
- Enjoyment (Q9): The Emotion technique stood out as the most enjoyable, with a mean score of 5.26.
- Notice Difference (Q10): User Selected was most recognized for distinct gameplay differences, with a score of 5.13.

To further explore player feedback from post-questionnaires, we conducted three distinct statistical analyses: the Friedman Test for identifying differences across multiple groups, the Mann–Whitney U Test for comparisons between two independent groups, and the Wilcoxon Signed-Rank Test for assessing differences between paired samples. The outcomes of these tests are detailed in Tables 11–14, highlighting how the different approaches to adjusting game difficulty shape aspects such as self-reported player engagement, fun, enjoyment, and excitement.

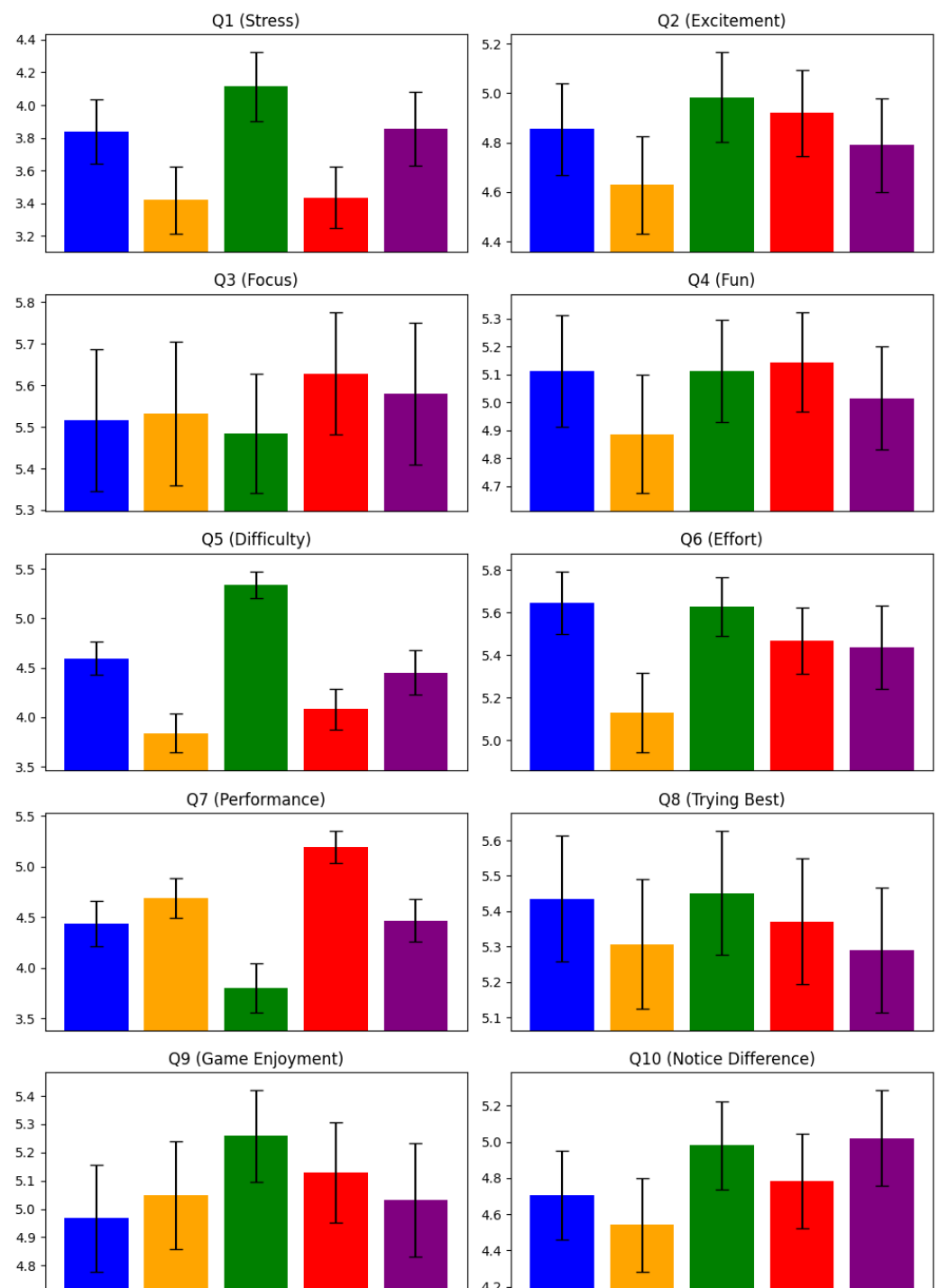


Figure 8. Mean results for all questions on post-questionnaire. Techniques are color-coded as follows: Combined (blue), Fixed (yellow), Emotion (green), Performance (red), and User Selected (purple).

The Friedman Test was employed to evaluate ordinal responses from the post-questionnaire, assessing differences in participant reactions across various testing conditions (see Table 11). This analysis revealed statistically significant differences in participant responses for several key areas:

- Stress (Q1): There were significant variations in stress levels reported by participants across the different DDA techniques, with a Friedman Test statistic of $\chi^2(4) = 23.135$. This suggests that some techniques are perceived as more stressful than others.

- Perceived Difficulty (Q5): Participants also perceived the difficulty of the techniques differently, as indicated by a Friedman Test statistic of $\chi^2(4) = 21.766$.
- Effort (Q6): The amount of effort required by different techniques varied significantly ($\chi^2(4) = 14.407$), indicating that some DDA settings demand more from players than others.
- Performance (Q7): Participants' self-assessed performance varied significantly across techniques ($\chi^2(4) = 21.322$).

Table 11. Post-questionnaire analysis using Friedman Test.

Questions	Friedman Test
Q1 (Stress)	$\chi^2(4) = 23.135, p < 0.050$
Q2 (Excitement)	$\chi^2(4) = 4.215, p = 0.378$
Q3 (Focus)	$\chi^2(4) = 2.675, p = 0.614$
Q4 (Fun)	$\chi^2(4) = 4.098, p = 0.393$
Q5 (Difficulty)	$\chi^2(4) = 21.766, p < 0.050$
Q6 (Effort)	$\chi^2(4) = 12.407, p < 0.050$
Q7 (Performance)	$\chi^2(4) = 21.322, p < 0.050$
Q8 (Trying Best)	$\chi^2(4) = 1.830, p = 0.767$
Q9 (Game Enjoyment)	$\chi^2(4) = 4.772, p = 0.311$
Q10 (Notice Difference)	$\chi^2(4) = 5.196, p = 0.268$

Table 12. Post-questionnaire analysis using Mann–Whitney U Test.

Questions	Combined	Fixed	Emotion	Performance	User Selected
Q1 (Stress)	U = 104.0, $p = 0.547$	U = 79.0, $p = 0.109$	U = 118.5, $p = 0.984$	U = 109.5, $p = 0.703$	U = 104.5, $p = 0.562$
Q2 (Excitement)	U = 87.0, $p = 0.199$	U = 100.0, $p = 0.444$	U = 88.0, $p = 0.213$	U = 80.5, $p = 0.123$	U = 81.0, $p = 0.128$
Q3 (Focus)	U = 102.0, $p = 0.494$	U = 96.0, $p = 0.354$	U = 74.0, $p = 0.068$	U = 97.0, $p = 0.375$	U = 93.0, $p = 0.297$
Q4 (Fun)	U = 108.5, $p = 0.673$	U = 94.0, $p = 0.316$	U = 86.5, $p = 0.191$	U = 107.0, $p = 0.630$	U = 96.5, $p = 0.367$
Q5 (Difficulty)	U = 111.0, $p = 0.746$	U = 49.0, $p < 0.050$	U = 92.5, $p = 0.282$	U = 103.5, $p = 0.535$	U = 112.0, $p = 0.779$
Q6 (Effort)	U = 93.5, $p = 0.298$	U = 112.0, $p = 0.778$	U = 68.0, $p < 0.050$	U = 102.0, $p = 0.491$	U = 101.5, $p = 0.482$
Q7 (Performance)	U = 102.0, $p = 0.497$	U = 54.5, $p < 0.050$	U = 84.0, $p = 0.161$	U = 82.0, $p = 0.133$	U = 109.5, $p = 0.703$
Q8 (Trying Best)	U = 99.5, $p = 0.433$	U = 115.5, $p = 0.888$	U = 84.5, $p = 0.167$	U = 73.0, $p = 0.064$	U = 113.0, $p = 0.809$
Q9 (Game Enjoyment)	U = 112.5, $p = 0.793$	U = 100.0, $p = 0.448$	U = 106.5, $p = 0.613$	U = 94.0, $p = 0.316$	U = 108.5, $p = 0.669$
Q10 (Notice Difference)	U = 114.0, $p = 0.841$	U = 107.0, $p = 0.631$	U = 108.5, $p = 0.675$	U = 97.0, $p = 0.378$	U = 118.5, $p = 0.984$

The Mann–Whitney U Test was employed to discern differences between in ordinal responses between the two gaming experience groups (casual vs. experienced) for each questionnaire item, detailed in Table 12. The analysis found no significant differences between the two groups across a range of questions, including Stress (Q1), Excitement (Q2), Focus (Q3), Fun (Q4), Trying Best (Q8), Game Enjoyment (Q9), and Noticeable Difference between techniques (Q10), indicating the uniformity in participants' perceptions across these aspects regardless of the DDA technique used. However, significant distinctions emerged between casual and experienced players in specific areas:

- Perceived Difficulty (Q5): A discernible difference was observed in the Fixed technique condition, where casual gamers perceived the game as more difficult than experienced gamers as indicated by the significant Mann–Whitney U value ($U = 49.0$).
- Effort (Q6): In regards to the Emotion technique, there was a significant difference in effort levels reported, with casual gamers experiencing and reporting higher levels of effort compared to experienced gamers ($U = 68.0$), suggesting it may elicit more engagement or challenge.
- Self-Rated Performance (Q7): The Fixed technique revealed a significant difference in how casual and experienced players rated their performance ($U = 54.5$), with experienced players feeling more effective or satisfied with their performance compared to casual players.

We utilized the Wilcoxon Signed-Rank Test for pairwise comparisons between different DDA techniques, abbreviated as Combined (CB), Fixed (FX), Emotion (EM), Performance

(PE), and User Selected (US). Results from this analysis are detailed in Tables 13 and 14, demonstrating significant findings:

- Stress (Q1): Significant differences were noted between the Emotion and Fixed techniques, with participants finding Emotion more stressful than Fixed ($Z = -3.404$).
- Perceived Difficulty (Q5): Participants rated Emotion as more challenging than Combined and Fixed, and Performance as significantly less difficult, highlighting the significant variance in perceived challenge (Z-scores: -3.357 , -4.057 , and -3.593 , respectively).
- Effort (Q6): Differences were significant between Combined and Fixed ($Z = -2.168$), and Fixed and Emotion ($Z = -2.995$), indicating that participants reported exerting more effort with Emotion compared to Fixed, and less effort with Fixed compared to Combined.
- Performance (Q7): Significant findings were noted for Combined and Performance ($Z = -2.818$), Fixed and Emotion ($Z = -3.337$), and Emotion and Performance ($Z = -4.089$) pairs. Participants felt they performed better with Performance compared to Combined and Emotion, and worse with Emotion compared to Fixed.

Table 13. Post-questionnaire analysis: pairwise comparisons using Wilcoxon Signed-Rank Test for questions 1–5.

Pairs	Q1	Q2	Q3	Q4	Q5
CB × FX	$Z = -2.168, p = 0.030$	$Z = -1.391, p = 0.164$	$Z = -0.041, p = 0.967$	$Z = -1.305, p = 0.192$	$Z = -2.437, p = 0.015$
CB × EM	$Z = -2.458, p = 0.014$	$Z = -0.724, p = 0.469$	$Z = -0.188, p = 0.851$	$Z = -0.053, p = 0.958$	$Z = -3.357, p < 0.005$
CB × PE	$Z = -2.277, p = 0.023$	$Z = -0.294, p = 0.769$	$Z = -1.027, p = 0.305$	$Z = -0.115, p = 0.909$	$Z = -1.875, p = 0.061$
CB × US	$Z = -0.407, p = 0.684$	$Z = -0.418, p = 0.676$	$Z = -0.515, p = 0.607$	$Z = -0.940, p = 0.347$	$Z = -0.161, p = 0.872$
FX × EM	$Z = -3.404, p < 0.005$	$Z = -1.990, p = 0.047$	$Z = -0.401, p = 0.688$	$Z = -1.093, p = 0.274$	$Z = -4.057, p < 0.005$
FX × PE	$Z = -0.165, p = 0.869$	$Z = -1.572, p = 0.116$	$Z = -0.808, p = 0.419$	$Z = -1.240, p = 0.215$	$Z = -0.653, p = 0.514$
FX × US	$Z = -2.136, p = 0.033$	$Z = -1.529, p = 0.126$	$Z = -0.606, p = 0.545$	$Z = -0.797, p = 0.425$	$Z = -2.099, p = 0.036$
EM × PE	$Z = -3.218, p < 0.005$	$Z = -0.320, p = 0.749$	$Z = -0.997, p = 0.319$	$Z = -0.369, p = 0.712$	$Z = -3.593, p < 0.005$
EM × US	$Z = -1.412, p = 0.158$	$Z = -0.843, p = 0.399$	$Z = -0.799, p = 0.424$	$Z = -0.739, p = 0.460$	$Z = -2.345, p = 0.019$
PE × US	$Z = -1.697, p = 0.090$	$Z = -0.856, p = 0.392$	$Z = -0.309, p = 0.757$	$Z = -0.615, p = 0.538$	$Z = -1.266, p = 0.206$

Table 14. Post-questionnaire analysis: pairwise comparisons using Wilcoxon Signed-Rank Test for questions 6–10.

Pairs	Q6	Q7	Q8	Q9	Q10
CB × FX	$Z = -3.220, p < 0.005$	$Z = -0.973, p = 0.330$	$Z = -0.602, p = 0.547$	$Z = -0.373, p = 0.709$	$Z = -0.386, p = 0.699$
CB × EM	$Z = -0.270, p = 0.787$	$Z = -2.401, p = 0.016$	$Z = -0.661, p = 0.509$	$Z = -2.399, p = 0.016$	$Z = -0.823, p = 0.410$
CB × PE	$Z = -1.149, p = 0.251$	$Z = -2.818, p < 0.005$	$Z = -0.288, p = 0.773$	$Z = -1.075, p = 0.282$	$Z = -0.638, p = 0.523$
CB × US	$Z = -0.892, p = 0.373$	$Z = -0.022, p = 0.983$	$Z = -1.304, p = 0.192$	$Z = -0.509, p = 0.610$	$Z = -1.716, p = 0.086$
FX × EM	$Z = -2.995, p < 0.005$	$Z = -3.337, p < 0.005$	$Z = -1.194, p = 0.232$	$Z = -1.384, p = 0.166$	$Z = -1.349, p = 0.177$
FX × PE	$Z = -1.810, p = 0.070$	$Z = -2.476, p = 0.013$	$Z = -0.230, p = 0.818$	$Z = -0.604, p = 0.546$	$Z = -0.945, p = 0.345$
FX × US	$Z = -2.099, p = 0.036$	$Z = -1.210, p = 0.226$	$Z = -0.044, p = 0.965$	$Z = -0.062, p = 0.951$	$Z = -1.725, p = 0.085$
EM × PE	$Z = -1.079, p = 0.281$	$Z = -4.089, p < 0.005$	$Z = -0.327, p = 0.744$	$Z = -0.710, p = 0.478$	$Z = -0.635, p = 0.525$
EM × US	$Z = -0.537, p = 0.591$	$Z = -2.181, p = 0.029$	$Z = -1.286, p = 0.198$	$Z = -1.266, p = 0.205$	$Z = -0.297, p = 0.766$
PE × US	$Z = -0.194, p = 0.847$	$Z = -2.452, p = 0.014$	$Z = -0.462, p = 0.644$	$Z = -0.686, p = 0.493$	$Z = -1.041, p = 0.298$

7. Discussion

Our study was a detailed exploration of DDA techniques in video games, analyzing three DDA approaches—performance-based, emotion-based, and a hybrid of the two—against two static difficulty settings (fixed and user-selected). Our objective was to better understand player preferences, the interplay between difficulty levels with player engagement and stress, and the impact of these difficulty adjustments on both performance metrics and physiological responses.

In regards to RQ1, which asks if participants prefer DDA over static difficulties in video games, the analysis yielded no significant differences in player preferences between DDA and static difficulties as indicated by both physiological sensor data and post-questionnaire feedback. The physiological metrics recorded did not demonstrate the expected variance, limiting our ability to draw definitive conclusions. Similarly, participant responses regarding excitement, fun, and overall game enjoyment did not reveal a statistically significant

preference for any approach to difficulty adjustments. The absence of significant findings for RQ1 suggests several possibilities. First, it may indicate that player preferences for DDA versus static difficulties are more intricate than initially hypothesized, possibly influenced by factors not captured in our study. Alternatively, the methodologies employed—both in terms of physiological measurement and self-reporting—may not have been sufficiently sensitive to detect subtle preferences or effects. This outcome emphasizes the complexity of measuring player engagement and satisfaction, highlighting the need for further research with refined methodologies or alternative metrics of player experience.

Similarly, our investigation into RQ2, which asks if participants have a preference between the three DDA techniques, did not yield significant results from either the physiological sensors or post-questionnaire analyses. This finding was unexpected, as we hypothesized there would be clear distinctions in player preferences based on the adaptive nature of each DDA approach. The lack of significant findings for RQ2 may reflect that participants do not have a strong preference for one DDA method over another, or that the differences between these techniques are not as perceptible to players as we anticipated as evidenced by Q10, where no significance was found, in the post-questionnaire which asks participants if they feel there is a noticeable difference between the current technique in comparison to those previously tested.

Investigating the efficacy of Flow Theory [7] in video games, RQ3 seeks to understand if higher difficulty levels, appropriately matched to player skill, lead to higher levels of stress and engagement. Our analysis, grounded in self-reported stress (Q1) and effort (Q6) from post-questionnaires, revealed significant correlations, particularly with the Emotion DDA technique, which presented the highest challenge. This technique's elevated difficulty—both perceived (reflected in Q5 responses) and measured in actual gameplay metrics—aligns with Flow Theory's principles, suggesting that an optimal challenge level can indeed foster engagement, despite increased stress.

The investigation into our final research question RQ4 revealed discernible differences in player performance metrics attributable to the deployment of diverse DDA techniques. Notably, metrics such as shot accuracy, damage dealt, survival time, and total scores exhibited significant statistical variances, underscoring the distinct effects of each DDA approach. Furthermore, other metrics, including damage taken, cows saved, and UFOs killed, while showing variance to a lesser extent, nonetheless contributed valuable insights into the complex effects of DDA on gameplay.

A pronounced disparity was observed in the difficulty level generated by different DDA techniques, which notably could have influenced many of the game performance metric findings. The Emotion technique, in particular, was characterized by a higher mean difficulty compared to other methods. This may have had a cascading effect that influenced the other game metrics. Players navigating the increased difficulty of the Emotion technique experienced larger amounts of damage taken, aligning with our observations of the technique's elevated challenge. Conversely, the Fixed technique, marked by its lower difficulty, led to reduced damage taken by players, highlighting the inverse relationship between difficulty level and player survivability.

Moreover, the analysis of secondary objectives, such as UFOs killed, further demonstrated the influence of DDA techniques on gameplay. The high difficulty associated with the Emotion technique resulted in fewer UFOs being eliminated, whereas the lower difficulty level of the Fixed technique facilitated greater success in this objective, reinforcing the hypothesized link between difficulty and objective completion. The survival time metrics further confirmed the significant role DDA techniques had in shaping the game experience, with the Emotion techniques having higher difficulty correlating with shorter player survival times, further illustrating the impact of the difficulty of providing sufficient challenge and ensuring player engagement and longevity.

Our findings contribute to the ongoing discussion on adaptive game design, suggesting complex player interactions with DDA that merit further exploration. The study's

limitations, especially in capturing the breadth of player experiences, set the stage for future research to more accurately tailor DDA techniques for improved player engagement.

8. Limitations

In our user study, we faced challenges with the devices used to gather physiological data. The EEG headset, used for measuring brain activity, often failed to connect reliably with participants with thick hair, leading to inconsistent signals. To preserve data integrity, we excluded the affected participants' data from our analysis. Additionally, the emotional variance captured by the device was lower than expected, which might be attributed to either the device's sensitivity or the game's inability to provoke strong emotional responses.

Our use of an optical heart rate monitor to gauge stress levels also fell short of expectations. Despite heart rate variability being a proven indicator of stress in prior studies, we did not observe significant differences during our tests. As a result, heart rate data were relegated to post-study analysis rather than being a primary measure of stress and engagement.

Furthermore, the study's participant diversity posed another limitation, especially regarding gender balance. With a participant pool of 27 males and only 4 females, there is a potential bias towards male gaming experiences and responses. This gender disparity limits the generalizability of our findings, as it may not fully account for gender-specific differences in stress response, engagement, and gaming behaviors. Future research should strive for more equitable gender representation to enhance the universality of the findings.

9. Conclusions

In this study, we explored the efficacy of Dynamic Difficulty Adjustment across a spectrum of player experiences in video games, contrasting their impact on both casual and experienced gamer groups. Despite the anticipation that DDA techniques would distinctly influence player preferences and emotional responses, our findings reveal a more complex interaction. No single DDA technique, including performance-based, emotion-based, or a hybrid approach, demonstrated clear superiority in enhancing the gaming experience or engaging players more deeply than static difficulty settings.

Our investigation into whether participants exhibited a preference for DDA over static difficulties or among the DDA techniques themselves did not yield significant preferences. This outcome suggests that player engagement with DDA is influenced by a broader range of factors than those captured within the scope of our study. Interestingly, the data did point to a potential alignment with Flow Theory, where the challenge provided by the Emotion DDA technique correlated with increased stress and effort, hinting at a sophisticated balance that might foster deeper engagement under certain conditions.

The modest effect of DDA on player performance metrics contrasted with its limited impact on emotional and physiological responses, illuminating the challenges in crafting adaptive difficulty settings that engage players both mechanically and emotionally. This distinction emphasizes the importance of further investigation into how DDA can be refined to suit a variety of player experiences, aiming to enhance gameplay in a way that complements both the emotional depth and physiological aspects of gaming.

Looking forward, improving our approach to emotion recognition becomes a key priority. The challenges faced with existing technologies, like the EEG device, highlight the need for more responsive and precise instruments to measure player emotions in real-time. Progress in this domain is essential for developing DDA strategies that are genuinely adaptive, potentially reshaping game design to provide experiences finely tuned to the player's emotional and cognitive states.

Our research adds to the evolving discussion on Dynamic Difficulty Adjustment, highlighting the complex interplay between game difficulty, player engagement, and emotional response. As we further examine and improve DDA techniques, our goal is to discover strategies that significantly better player experiences, thereby making games more inclusive and enjoyable for all.

In the future, we would like to investigate more into enhancing the Emotion-based DDA technique. This involves increasing the precision with which we assess players' emotional states and refining our methods for adjusting difficulty based on these measurements. Additionally, there is potential for incorporating machine learning algorithms to fine-tune game difficulties, specifically tailored to accommodate a wider range of player abilities—from complete novices to highly experienced gamers. Moreover, applying DDA techniques to a broader range of game genres beyond the first-person shooter studied here could provide insightful findings into the versatility and effectiveness of these strategies across different gaming contexts.

Author Contributions: Conceptualization, N.F.; Methodology, N.F. and A.K.K.; Software, N.F.; Formal analysis, N.F.; Investigation, N.F.; Resources, A.K.K.; Data curation, N.F.; Writing—original draft, N.F.; Writing—review & editing, N.F. and A.K.K.; Visualization, N.F.; Supervision, A.K.K.; Project administration, A.K.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted per the Declaration of Helsinki and approved by the Institutional Review Board (IRB) of the University of Louisiana at Lafayette. (protocol code FA20-34CMCP and approved on 15/12/2020).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. The data are not publicly available due to the ethical approval of having participants' informed consent, which included participants being informed that we will remove all personally identifiable information before sharing data with universities and/or research institutions.

Acknowledgments: We would like to acknowledge all members of the Human-Computer Interaction (HCI) Laboratory for useful feedback during the planning phases of this project. This research used equipment purchased with support from the Louisiana Board of Regents under contract No. LEQSF(2022-25)-RD-A-24.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DDA	Dynamic Difficulty Adjustment
FPS	First-Person-Shooter
MMO	Massively-Multiplayer Online
MOBA	Multiplayer Online Battle Arena
EEG	Electroencephalogram
fNIRS	Functional near-infrared spectroscopy
GSR	Galvanic skin response
HRV	Heart rate variability
UFO	Unidentified flying object
ADS	Aim down sight
PC	Personal computer
MANOVA	Multivariate analysis of variance
CB	Combined
FX	Fixed
EM	Emotion
PE	Performance
US	User Selected

References

1. Jennett, C.; Cox, A.L.; Cairns, P.; Dhoparee, S.; Epps, A.; Tijs, T.; Walton, A. Measuring and defining the experience of immersion in games. *Int. J. Hum.-Comput. Stud.* **2008**, *66*, 641–661. [[CrossRef](#)]

2. Zohaib, M. Dynamic Difficulty Adjustment (DDA) in Computer Games: A Review. *Adv. Hum.-Comput. Interact.* **2018**, *2018*, 1–12. [CrossRef]
3. Kim, H.G.; Cheon, E.J.; Bai, D.S.; Lee, Y.H.; Koo, B.H. Stress and Heart Rate Variability: A Meta-Analysis and Review of the Literature. *Psychiatry Investig.* **2018**, *15*, 235–245. [CrossRef]
4. Taelman, J.; Vandeput, S.; Spaepen, A.; Van Huffel, S. Influence of Mental Stress on Heart Rate and Heart Rate Variability. In Proceedings of the 4th European Conference of the International Federation for Medical and Biological Engineering, Antwerp, Belgium, 23–27 November 2008; Magjarevic, R., Nagel, J.H., Vander Sloten, J., Verdonck, P., Nyssen, M., Haueisen, J., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; Volume 22, pp. 1366–1369. [CrossRef]
5. Salahuddin, L.; Cho, J.; Jeong, M.G.; Kim, D. Ultra Short Term Analysis of Heart Rate Variability for Monitoring Mental Stress in Mobile Settings. In Proceedings of the 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Lyon, France, 22–26 August 2007; pp. 4656–4659. [CrossRef]
6. Bicalho, L.F.; Baffa, A.; Feijó, B. A Dynamic Difficulty Adjustment Algorithm with Generic Player Behavior Classification Unity Plugin In Single Player Games. In Proceedings of the 22nd Brazilian Symposium on Games and Digital Entertainment, New York, NY, USA, 6–9 November 2024; SBGames '23; pp. 76–85. [CrossRef]
7. Csikszentmihalyi, M. *Flow: The Psychology of Optimal Experience*; Harper: Manhattan, NY, USA, 1990.
8. Bowman, N.D.; Tamborini, R. Task demand and mood repair: The intervention potential of computer games. *New Media Soc.* **2012**, *14*, 1339–1357. [CrossRef]
9. Lemmens, J.S.; von Münchhausen, C.F. Let the beat flow: How game difficulty in virtual reality affects flow. *Acta Psychol.* **2023**, *232*, 103812. [CrossRef] [PubMed]
10. Hunnicke, R.; Chapman, V. AI for dynamic difficulty adjustment in games. In Proceedings of the Challenges in Game Artificial Intelligence Workshop, Advancement of Artificial Intelligence (AAAI), San Jose, CA, USA, 25–29 July 2004; Volume 2.
11. Hunnicke, R. The case for dynamic difficulty adjustment in games. In Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology, Valencia, Spain, 15–17 June 2005; pp. 429–433. [CrossRef]
12. Hristova, D. Dynamic Difficulty Adjustment (DDA) in First Person Shooter (FPS) Games. Available online: https://www.academia.edu/download/32184026/HristovaDayana_DynamicDifficultyAdjustment.pdf (accessed on 4 April 2023).
13. Pfau, J.; Smeddinck, J.D.; Malaka, R. 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, Honolulu, HI, USA, 25–30 April 2020; pp. 1–10. [CrossRef]
14. Silva, M.P.; Silva, V.d.N.; Chaimowicz, L. Dynamic Difficulty Adjustment on MOBA Games. *arXiv* **2017**, arXiv:1706.02796. <https://doi.org/10.48550/ARXIV.1706.02796>.
15. Romero-Mendez, E.A.; Santana-Mancilla, P.C.; Garcia-Ruiz, M.; Montesinos-López, O.A.; Anido-Rifón, L.E. The Use of Deep Learning to Improve Player Engagement in a Video Game through a Dynamic Difficulty Adjustment Based on Skills Classification. *Appl. Sci.* **2023**, *13*, 8249. [CrossRef]
16. Huber, T.; Mertes, S.; Rangelova, S.; Flutura, S.; Andre, E. Dynamic Difficulty Adjustment in Virtual Reality Exergames through Experience-driven Procedural Content Generation. In Proceedings of the 2021 IEEE Symposium Series on Computational Intelligence (SSCI), Orlando, FL, USA, 5–7 December 2021. [CrossRef]
17. Francillette, Y.; Tremblay, H.; Bouchard, B.; Lescieux, S.; Rozon, M.; Linard, J. Automated Difficulty Assessment Model for Platformer Games: A Comprehensive Approach. In Proceedings of the 2023 IEEE Gaming, Entertainment, and Media Conference (GEM), St. Michael, Barbados, 27–30 November 2023; pp. 1–6. [CrossRef]
18. Liu, C.; Agrawal, P.; Sarkar, N.; Chen, S. Dynamic Difficulty Adjustment in Computer Games Through Real-Time Anxiety-Based Affective Feedback. *Int. J. Hum.-Comput. Interact.* **2009**, *25*, 506–529. [CrossRef]
19. Park, S.; Sim, H.; Lee, W. Dynamic Game Difficulty Control by Using EEG-based Emotion Recognition. *Int. J. Control. Autom.* **2014**, *7*, 267–272. [CrossRef]
20. Fernandez B., H.D.; Mikami, K.; Kondo, K. Adaptable Game Experience Based on Player's Performance and EEG. In Proceedings of the 2017 Nicograph International (NicoInt), Kyoto, Japan, 2–3 June 2017; pp. 1–8. [CrossRef]
21. Stein, A.; Yotam, Y.; Puzis, R.; Shani, G.; Taieb-Maimon, M. EEG-triggered dynamic difficulty adjustment for multiplayer games. *Entertain. Comput.* **2018**, *25*, 14–25. [CrossRef]
22. Afergan, D.; Peck, E.M.; Solovey, E.T.; Jenkins, A.; Hincks, S.W.; Brown, E.T.; Chang, R.; Jacob, R.J. Dynamic difficulty using brain metrics of workload. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA, 18–23 April 2014; CHI '14; pp. 3797–3806. [CrossRef]
23. Moschovitis, P.; Denisova, A. Keep Calm and Aim for the Head: Biofeedback-Controlled Dynamic Difficulty Adjustment in a Horror Game. *IEEE Trans. Games* **2022**, *15*, 386–377. [CrossRef]
24. Rosa, M.P.C.; Dos Santos, E.A.; De Moraes, I.L.R.; E Silva, T.B.P.; Sarmet, M.M.; Castanho, C.D.; Jacobi, R.P. Dynamic Difficulty Adjustment Using Performance and Affective Data in a Platform Game. In *HCI International 2021—Late Breaking Papers: Design and User Experience*; Stephanidis, C., Soares, M.M., Rosenzweig, E., Marcus, A., Yamamoto, S., Mori, H., Rau, P.L.P., Meiselwitz, G., Fang, X., Moallem, A., Eds.; Springer International Publishing: Cham, Switzerland, 2021; Volume 13094, pp. 367–386. [CrossRef]
25. Nguyen, D. Improving Game Experience Using Dynamic Difficulty Adjustment Based on Physiological Signals. Master's Thesis, Technical University of Denmark, DTU Compute, Kongens Lyngby, Denmark, 2014.

26. Karavidas, L.; Apostolidis, H.; Tsiatsos, T. Usability Evaluation of an Adaptive Serious Game Prototype Based on Affective Feedback. *Information* **2022**, *13*, 425. [[CrossRef](#)]
27. Guo, Z.; Thawonmas, R.; Ren, X. Rethinking dynamic difficulty adjustment for video game design. *Entertain. Comput.* **2024**, *50*, 100663. [[CrossRef](#)]
28. Koskinen, A.; McMullen, J.; Hannula-Sormunen, M.; Ninaus, M.; Kiili, K. The strength and direction of the difficulty adaptation affect situational interest in game-based learning. *Comput. Educ.* **2023**, *194*, 104694. [[CrossRef](#)]
29. Darzi, A.; McCrea, S.M.; Novak, D. User Experience with Dynamic Difficulty Adjustment Methods for an Affective Exergame: Comparative Laboratory-Based Study. *JMIR Serious Games* **2021**, *9*, e25771. [[CrossRef](#)] [[PubMed](#)]
30. Emotiv. Performance Metrics. 2024. Available online: <https://www.emotiv.com/tools/knowledge-base/performance-metrics/performance-metrics/> (accessed on 16 April 2024).
31. Holm, S. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* **1979**, *6*, 65–70.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.