Proactive or Reactive? Comparison of Dynamic Difficulty Adjustment Effects on Player Experience and Physical Effort in a Running Exergame

[1st] Sittikrai Chorrojprasert , [2nd] Ben Kirman, [3rd] Myat Thura Aung, and [4th] Alena Denisova
[1st] School of IT, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand (sittikraichor.contact@gmail.com)
[2nd] [4th] University of York, York, UK ({ben.kirman, alena.denisova}@york.ac.uk)
[3rd] Independent researcher, York, UK (aungmthura@gmail.com)

***Abstract*—Maintaining engagement in exergames should balance game challenge with players' abilities. This study investigates how reactive and proactive dynamic difficulty adjustment (DDA) functions shape player experience (PX) and physical effort in the custom-built running exergame, *"ExerWitch"*. While reactive DDA adapts difficulty only after failure thresholds are met, proactive DDA continuously monitors performance to adjust parameters before setback occurrence. In a within-subjects study ($N = 36$), we compared these approaches using heart-rate based exertion metrics and standardised PX indicators. Our results indicate that compared to reactive DDA, proactive DDA significantly enhanced challenge-skill balance ($p = .017, r = .40$), increased perceived competence ($p = .002, r = .51$), and reduced competence frustration ($p = .040, r = .34$). No significant differences were observed in perceived exertion or physical effort. These findings suggest proactive DDA better supports PX in locomotion-based exergames with repeated failure penalties, offering empirical insights into how adaptation timing influences exertion-based gameplay.**

***Index Terms*—exergame, video game, exertion, dynamic difficulty adjustment, DDA, player experience, physical effort**

I. Introduction

Physical activity has been integrated into commercial video games for several decades [1]. Within this domain, exergames—or exertion games—leverage physical effort as a core gameplay mechanic to enhance engagement beyond traditional fitness activities [2, 3]. By translating bodily movements into virtual interactions via motion-sensing technologies, these systems utilise the motivational affordances of digital play [4] to promote sustained energy expenditure in an interactive manner.

A central objective in exergame design is maintaining flow, defined as the balance between game challenge and player skill [5]. Prior work [6, 7] highlights the importance of this balance: excessive difficulty may lead to frustration and physical exhaustion, whereas insufficient challenge can reduce engagement and limit exercise benefits [8]. To address this, dynamic difficulty adjustment (DDA) has been widely adopted to adapt gameplay parameters in real time based on player performance or physiological feedback [9, 10]. Compared with manual adjustment, DDA enables seamless adaptation without interrupting gameplay [11].

However, most exergame research—including prior studies [12–14]—focuses on reactive DDA ($DDA_{Reactive}$). These approaches infer the need for adaptation from observed performance breakdowns or failure-related indicators, adjusting difficulty only after players begin to struggle. While effective in traditional digital games [15], this paradigm assumes that failure primarily functions as informative feedback, allowing players to reassess strategies without significant consequence.

In exergames, this assumption may not hold. Because physical exertion constitutes the primary challenge, failure often conflicts with gameplay flow and imposes an additional physical demand [16]. Events such as enemy collisions, forced restarts, or "Game Over" states can interrupt exercise movement and undermine perceived competence [12]. When repeated, these interruptions may increase frustration and negatively affect player experience (PX) [17, 18]. Consequently, DDA approaches that rely on failure-driven signals risk undermining both gameplay continuity and sustained physical activity [12].

To address this limitation, we propose proactive DDA ($DDA_{Proactive}$), a predictive approach that anticipates player struggle and adjusts difficulty before failure occurs. By allowing for smoother alignment between physical effort and game demands, this approach aims to preserve positive PX without interruption of movement from dealing with failures.

Despite its potential, proactive DDA remains under-explored in exergame contexts. This study addresses this gap by comparing $DDA_{Proactive}$ and $DDA_{Reactive}$ in a custom running exergame, ExerWitch. We examine their effects on PX and physical effort, guided by the following research question: How do $DDA_{Reactive}$ and $DDA_{Proactive}$ influence PX and physical effort in an exergame context?

II. Background

In everyday contexts, failure is typically avoided; however, video games provide a safe environment in which players actively seek challenges that involve loss or avatar death [19]. This phenomenon, often referred to as the "paradox of failure" [20], suggests that players willingly engage with

979-8-3315-9476-3/26/$31.00 ©2026 IEEE

unpleasant failure states to satisfy psychological needs, such as the experience of competence gained through overcoming obstacles [17, 21].

Within game research, the terms "setback" and "failure" are frequently used interchangeably, despite representing functionally distinct concepts. A setback is generally defined as an event that delays or hinders progress. In games, setbacks often take the form of punishments that require players to invest additional time or resources to resume progress [22]. While Juul [23] conceptualises all failures as a subset of setbacks, other researchers [12, 17, 24, 25] argue the inverse relationship. Specifically, Frommel et al. [17], Cuerdo et al. [24], Aytemiz and Smith [25] describe setbacks as a subset of failure—namely, temporary obstacles that delay progress while offering immediate opportunities for players to refine skills or adapt strategies at relatively low cost. In contrast, failure is characterised by the perception that a goal is unattainable or requires substantial effort to recover.

Importantly, the perception of failure is also shaped by player-defined goals, which may diverge from those intended by the game designer [17, 26]. For example, a "Game Over" screen may be interpreted as a mere setback if it does not conflict with the player's personal objectives [25]. Accordingly, we define failure as the inability to achieve a perceived goal that remains mechanically attainable within the rules of the game.

Failure and setbacks in games are not always negative; they are often necessary for positive PX [17]. Overcoming difficult rules or time limits can increase a player's sense of accomplishment [27]. Players often seek a "well-played game" that balances difficulty, rather than one that is too easy [26]. Success feels more rewarding when failure is a real possibility [23, 28].

While high challenge can cause temporary anxiety, it often leads to positive feelings once the player succeeds [19]. Research suggests that a harder struggle can lead to greater satisfaction upon winning [29]. However, if failure happens too often, it leads to frustration and may cause the player to stop playing entirely [17, 18, 30, 31].

In exergames, failure is more complex because it involves physical exertion. Unlike traditional games where a player can rest and retry, exergame players face physical fatigue. This exhaustion makes it harder to persist through repeated failures [24]. This dynamic is similar to sports, where failure can cause stress and frustration, but overcoming it builds resilience and better strategy [32]. Because exergames require both mental and physical effort, finding the right balance of challenge is critical to prevent players from quitting due to exhaustion.

As failures originating from poorly calibrated challenges can degrade PX, DDA functions are commonly used to mitigate the gap between a player's ability and the game's difficulty level [33, 34]. A systematic review by Seyderhelm and Blackmore [35] demonstrates that DDA is applied across a wide spectrum of genres—from casual titles like Tetris and Pong to complex multiplayer role-playing games and first-person shooters. These functions typically monitor performance or physiological data to modify game parameters such as time constraints, task complexity, and speed in real-time.

DDA design is fundamentally rooted in Flow Theory [8]—maintaining the balance between challenge and skill—and Self-Determination Theory (SDT) [21]—promoting competence to enhance engagement [35]. Several studies confirm that flow-based DDA enhances the challenge-skill balance. Ang [36] found that performance-based DDA in Tetris improved this balance. Also, Climent et al. [37] and Lora-Ariza et al. [38] demonstrated that reinforcement-learning-based DDA significantly increased flow experiences and win rates, particularly for novice players.

In the context of exergames, DDA is primarily used to balance game demands with both player ability and physical fitness [5, 6]. In low-intensity exergames, such as the Pong variations studied by Goršič et al. [33] and Ajani and Mallipeddi [39], the DDA mechanism is ruled primarily by outcome-based failure metrics. Specifically, adjustments to ball speed and paddle size are triggered by changes in the relative score—a metric that essentially counts the frequency of a player's failure to return the ball. Similarly, the multi-modal system proposed by Darzi et al. [14] utilises performance data as the primary classifier, meaning the DDA only intervenes once the player's error rate indicates an existing mismatch between skill and challenge.

In more complex, high-intensity exergames, DDA functions often use threshold-based failure metrics to manage physical exertion. For instance, in Plunder Planet [7] and ExerCube [40], the "performance score" used for difficulty adjustment is often a direct reflection of the player's mistakes—such as missing a gate or failing to avoid an obstacle. These games monitor for a specific number of "in-game errors" before lowering the difficulty.

The potential limitation of these works is their reactive timing. Because the adjustment logic depends on failure-metrics, the game must allow the player to experience a set number of setbacks before it can act. While some literature [18, 19, 23, 28, 29] suggests that overcoming these failure-driven challenges can enhance a sense of mastery and competence, others [12, 17, 18, 30, 31] argue that reactive loops may lead to negative PX, such as increased frustration and reduced competence satisfaction.

Despite the potential for negative affect in reactive functions, proactive DDA—which aims to prevent failure by adjusting difficulty before a setback occurs—is rarely adopted in exergame research. While a proactive function may mitigate frustration, it also risks reducing the intense experience that stems from the failure-recovery cycle, potentially lowering the player's overall effort, as certain players rely on positive or motivational frustration to stimulate persistence and heighten satisfaction upon subsequent mastery [17, 41]. Due to the lack of empirical studies on this distinction, it remains unclear how proactive versus reactive timing influences players in the exergame context.

This paper aims to fill this gap by providing empirical data on how these two approaches impact PX and physical

effort in a running exergame. Our findings intend to provide design guidelines for future DDA functions that promote both engagement and physical exertion.

III. Hypotheses

Based on the literature review, the following dependent variables—potentially influenced by changes in game challenge levels—were identified, and corresponding hypotheses were formulated to address the research question through statistical testing. However, due to the limited number of prior studies using $DDA_{Proactive}$ and $DDA_{Reactive}$, all hypotheses were stated as non-directional to avoid bias and allow for the detection of effects in either direction

While PX is a multidimensional construct, this study focuses specifically on competence satisfaction, competence frustration, and challenge–skill balance. This targeted scope aligns with SDT and Flow Theory, which are related to DDA.

- H_1: Average heart rate differs between running exergames with $DDA_{Proactive}$ and $DDA_{Reactive}$.
- H_2: Perceived exertion differs between running exergames with $DDA_{Proactive}$ and $DDA_{Reactive}$.
- H_3: Competence satisfaction ($H_{3.A}$) and competence frustration ($H_{3.B}$) differ between running exergames with $DDA_{Proactive}$ and $DDA_{Reactive}$.
- H_4: Challenge-skill balance differs between running exergames with $DDA_{Proactive}$ and $DDA_{Reactive}$.

IV. Running Exergame: "ExerWitch"

This study used *ExerWitch* (Figure 1), a custom indoor running exergame developed in Unity 2021.3.8f1, to evaluate both DDA functions. Running in place was selected as the core mechanic because its wide intensity range—spanning from light to vigorous exertion [42]—allows DDA algorithms to adjust challenge levels across a broad, physically perceptible spectrum. This design purposefully avoids the "floor effect" common in low-intensity exergames like *Pong* [14, 39], enabling players to smoothly modulate game difficulty by altering their running velocity.



Fig. 1: A screenshot from the *ExerWitch* game of the witch running down a path away from the dragon.



Fig. 2: Study room setup and a participant playing *ExerWitch* by running in place and using arm movements to steer the avatar.

The game features an endless runner theme similar to *Temple Run* or *Subway Surfers*, allowing participants to rapidly familiarise themselves with the objectives. Players control a virtual witch fleeing a chasing dragon by performing continuous stationary running in front of a display. To isolate the physical challenge to velocity modulation, no secondary obstacles (e.g. jumping) were included. Failing to maintain sufficient speed triggers a dragon attack (*setback: collision*), causing temporary immobility, increased completion time, and energy depletion (*punishment*). If the energy metric reaches zero, the session terminates (*failure*).

The use of a temporary mobility penalty, rather than score or health-point reduction, was an intentional design choice. Less disruptive penalties may reduce the experiential distinction between reactive and proactive adaptation, whereas halting avatar movement creates a more immediate interruption to gameplay flow and physical momentum.

Participants wore a wrist-worn smartwatch (Samsung Galaxy Watch 4) for heart rate monitoring, alongside an ankle-worn sensor array for step frequency and intensity tracking. Real-time posture detection was captured via a tripod-mounted smartphone using Google's ML Kit framework to monitor upper-body joints. Data streams were transmitted via Bluetooth to the game engine and mirrored onto a television display 1.5 meters ahead (Figure 2). Prior to the formal evaluation, the game and its baseline DDA configurations were fine-tuned through a two-iteration pilot phase with four distinct participants (two per iteration) who did not take part in the main study.

V. Dynamic Difficulty Adjustment Functions

The DDA functions in *ExerWitch* are theoretically grounded in Flow theory [5, 8], aiming to maintain an equilibrium between the game's physical demands and the player's exertion capacity. In this context, *challenge* is defined as the physical requirement to maintain a specific speed to avoid the chasing dragon, while *skill* refers to the player's ability to regulate their running speed.

Both DDA conditions dynamically manipulate the dragon's movement speed (V_{dragon}) to calibrate difficulty, but differ fundamentally in their triggering criteria. We defined d_{diff} as the spatial distance (in game world units) between the player's avatar and the dragon. This variable serves as a predictive proxy for player performance; a decreasing d_{diff} indicates that the player's physical output is insufficient to maintain the running speed required to evade capture.

Also, the frequency of setbacks—defined as discrete collision events between the dragon and the avatar—serves as a retrospective measure of player performance, capturing instances where their performance failed to meet the game's challenge level.

A. Common Control Logic

To ensure the experience remained engaging and followed established game design principles, both DDA conditions shared a common baseline logic:

- **Proximity Maintenance:** To ensure a consistent challenge (the "rubber-band effect" [43]), if the player's

avatar creates a significant lead ($d_{diff} > 15$), the game moves the dragon closer to prevent the challenge from becoming trivial.
- **Success Reward:** If a player successfully evades the dragon within the proximity zone ($2 \leq d_{diff} \leq 4$) for 15 consecutive seconds, V_{dragon} is reset to its initial value (V_{init}), rewarding the player's sustained effort with an easier gameplay phase.
- **Difficulty Scaling:** Every 10 seconds, if no setbacks have occurred, the game executes a probabilistic scaling check ($P_{inc} = 0.6$). If successful, V_{dragon} increases to progress the challenge to improve player performance.

B. Experimental Conditions: Reactive vs. Proactive

The core independent variable of this study is the timing of the adaptation trigger. The two conditions function as follows:

1) Reactive DDA ($DDA_{Reactive}$): The DDA_{Reactive} condition functions as an event-driven corrective loop. It relies on explicit failure as the primary data point for difficulty calibration.

- **Trigger:** A "setback" event (collision between dragon and avatar).
- **Adjustment Logic:** The DDA remains at the current difficulty level until a setback threshold ($\eta \geq 2$) is reached. Upon the second collision, the game executes a hard reset, moving the dragon backward and resetting V_{dragon} to V_{init}. This allows the player to recover, but only after the exercise rhythm has been interrupted, as the avatar is temporarily unable to move as a form of punishment.

2) Proactive DDA ($DDA_{Proactive}$): In contrast, $DDA_{\text{Proactive}}$ acts as a threshold-driven anticipatory regulator. It focuses on preventing performance breakdown before it occurs.

- **Trigger:** Spatial proximity threshold ($d_{diff} = 2$).
- **Adjustment Logic:** The DDA continuously monitors the distance between the player and the dragon. When the dragon enters the "Critical Zone" ($d_{diff} \leq 2$), the DDA triggers an anticipatory reset of V_{dragon} and moves the dragon backward.

By intervening before a collision occurs, this condition prevents the "additional physical demand" of a physical restart—minimising the player's awareness of the adaptive assistance while preserving the continuity of their running.

The specific thresholds for above variables were fine-tuned during pilot testing to ensure effective difficulty scaling within the 10-minute experimental sessions. Table I shows the summary of both DDA functions.

TABLE I: Comparison of DDA_{Reactive} and $DDA_{\text{Proactive}}$

Aspect	DDA_{Reactive}	$DDA_{\text{Proactive}}$
Trigger	Setback ($\eta \geq 2$ collisions)	Proximity ($d_{\text{diff}} \leq 2$)
Timing	Post-failure	Pre-failure
Impact	Punishment (forced avatar stop)	No punishment
Goal	Failure-driven	Anticipatory

VI. METHODOLOGY

A. Experimental Design

The study was approved by the departmental physical activity committee and adopted a within-subjects design with randomised counterbalancing to mitigate order effects. Participants attended a single 1-hour session consisting of two 10-minute exergame segments, each featuring a different DDA condition (DDA_{Reactive} and $DDA_{\text{Proactive}}$). The session commenced with an intake briefing and informed consent collection, followed by a tutorial segment to familiarise participants with the game. To ensure data consistency and eliminate carry-over effects, a 5-minute rest period followed the tutorial to allow participants' heart rates to return to baseline, at which point resting heart rate (RHR) was recorded.

At the start of each experimental condition, participants were reminded of the primary game goals, core challenges, and failure consequences. This instruction was explicitly emphasised to prevent unintended player-defined goals or mismatched expectations [17, 26]. Following each 10-minute gameplay period, participants completed post-session questionnaires assessing PX. A mandatory 10-minute rest period was enforced between conditions to mitigate physiological fatigue. RHR was measured again at the conclusion of this rest period—before the second game session—to confirm that no significant differences existed in baseline heart rate. This step was used to mitigate order effect concerns regarding average heart rate (AHR) collected during active playtime.

To ensure sample adequacy, an a priori power analysis was conducted targeting 80% power at an alpha level of .05, with effect size estimates derived from related exergame literature [14, 33, 40, 44]. A total of 36 participants completed the study (18 male, 18 female; mean age = 28 years, $SD = 7.18$).

B. Measure and Analysis

The following observed variables were measured to test the hypotheses and address the research question. Heart rates (beats per minute) were recorded using a Samsung Galaxy Watch 4, and perceived exertion was rated on Borg's Rating of Perceived Exertion (BRPE) scale [45].

PX was operationalised using two validated psychometric instruments. Competence satisfaction and frustration were measured via the Basic Needs in Games Scale (BANGS) [46], using six items each on a 7-point Likert scale. Challenge-skill balance was examined using four items from the Flow State Scale (FSS) [47].

IBM SPSS 28.0.1.1 was used for data analysis and hypothesis testing. Data preprocessing included normality checks and outlier detection. A Wilcoxon signed-rank test was applied for non-normally distributed measures, while a paired-samples t-test was used for AHR, assuming a normal distribution. The alpha level was set at .05.

VII. Results

Descriptive statistics and t-test results comparing $\text{DDA}_{Reactive}$ and $\text{DDA}_{Proactive}$ are presented in Table II, while descriptive statistics, internal consistency scores, and Wilcoxon signed-rank test results for non-parametric data are shown in Table III.

A. Average Heart Rate

To assess order effects on AHR, RHR measured before each session were compared. No significant difference was found, $t(35) = .766$, $p = .449$, Cohen's $d = .13$ (negligible effect), suggesting no carry-over effect at the start of gameplay.

The analysis compared AHR during the playtime sessions to test the hypothesis H_1. Due to an application failure in the heart rate recorder, the AHR of 32 out of 36 participants were recorded and used for analysis.

The results indicated that the AHR in $\text{DDA}_{Reactive}$ ($M \pm S.D. = 148.85 \pm 16.43$) did not significantly differ from the AHR in $\text{DDA}_{Proactive}$ ($M \pm S.D. = 146.13 \pm 17.64$), $t(31) = 1.144$, $p = .261$, Cohen's $d = .20$ (small effect). Therefore, the null hypothesis of H_1 was not rejected.

B. BRPE: Perceived Exertion

The analysis revealed no significant difference in perceived exertion between the conditions, $Z = -.489$, $N = 36$, $p = .625$, $r = .08$ (negligible effect). Therefore, the null hypothesis of H_2 was not rejected.

C. BANGS: Competence Satisfaction and Frustration

The analysis showed that playing the exergame with $\text{DDA}_{Proactive}$ ($M \pm S.D. = 6.07 \pm .81$, *Mdn* = 6.33, *IQR* = 1.25) resulted in higher competence satisfaction compared to $\text{DDA}_{Reactive}$ ($M \pm S.D. = 5.28 \pm 1.31$, *Mdn* = 5.17, *IQR* = 2.26), $Z = -3.048$, $N = 36$, $p = .002$, $r = .51$ (large effect).

Additionally, the analysis revealed that playing the exergame with $\text{DDA}_{Proactive}$ ($M \pm S.D. = 2.51 \pm 1.17$, *Mdn* = 2.50, *IQR* = 2.26) resulted in less competence frustration compared to $\text{DDA}_{Reactive}$ ($M \pm S.D. = 3.09 \pm 1.61$, *Mdn* = 3.00, *IQR* = 2.75), $Z = -2.052$, $N = 36$, $p = .040$, $r = .34$ (medium effect). Therefore, the null hypotheses of $H_{3.A}$ and $H_{3.B}$ were rejected.

Figure 3 and Figure 4 present box plots illustrating the distribution of BANGS: Competence Satisfaction and BANGS: Competence Frustration, respectively, under the $\text{DDA}_{Reactive}$ and $\text{DDA}_{Proactive}$ conditions.

D. FSS: Challenge-skill Balance

The analysis revealed a significant difference in challenge-skill balance between the conditions, $Z = -2.388$, $N = 36$, $p = .017$, $r = .40$ (medium effect). Playing the exergame with $\text{DDA}_{Proactive}$ ($M \pm S.D. = 5.57 \pm .97$, *Mdn* = 5.63, *IQR* = 1.69) was perceived as having a higher challenge-skill balance compared to playing with $\text{DDA}_{Reactive}$ ($M \pm S.D. = 5.15 \pm 1.00$, *Mdn* = 5.13, *IQR* = 1.50). Consequently, the null hypothesis of H_4 was rejected.

The distribution of FSS: Challenge-Skill Balance in the $\text{DDA}_{Reactive}$ and $\text{DDA}_{Proactive}$ conditions is shown in box plots in Figure 5.

VIII. Discussion

This study investigated how DDA_{Reactive} and $DDA_{\text{Proactive}}$ affect PX and physical effort in a running exergame. Both approaches aim to maintain an optimal challenge–skill balance to support engagement [8]. However, our findings suggest that the timing of difficulty adaptation may play an important role in shaping PX during exertion-based gameplay.

A key contribution of this work is the significant increase in competence satisfaction ($p = .002$, $r = .51$) observed under $DDA_{\text{Proactive}}$. Previous DDA research (e.g. [7, 33, 39, 40]) has primarily focused on reactive adaptation mechanisms that adjust difficulty after players encounter observable setbacks. While such approaches may function effectively in cognitively or emotionally driven gameplay contexts [15], they also assume that failure primarily serves as informational feedback that enables players to reassess strategies and adapt their behaviour.

In our exergame, however, failure leads to the interruption of physical movement, creating a psychological conflict when the avatar is halted while the player's physical body remains active. Because physical exertion is the primary interaction loop, recovering from setbacks requires additional physical demand. Previous work suggests that repeated failures before a DDA_{Reactive} intervention occurs contribute to negative PX [12]. Our results are consistent with the interpretation that proactive adaptation reduces these negative experiential consequences by regulating difficulty before performance breakdown occurs.

In the implemented $DDA_{\text{Proactive}}$, the game subtly repositioned the dragon before attacks occurred, reducing the likelihood of repeated collisions while minimising player awareness of external assistance. This approach helped preserve continuous gameplay flow and a stable sense of competence. These findings align with Self-Determination Theory (SDT) [21], which suggests that competence satisfaction emerges when individuals feel effective in achieving desired outcomes. By reducing repeated breakdowns in performance, $DDA_{\text{Proactive}}$ may support players' perceived ability to manage gameplay challenges without making adaptive assistance overly salient.

The improvement in challenge–skill balance ($p = .017$, $r = .40$) further highlights the importance of adaptation timing in exergames. While classical flow models emphasise balancing challenge relative to player ability [5], our findings suggest that when difficulty adjustment occurs may also influence PX. Difficulty regulation implemented before explicit performance breakdown may better preserve the perception of equilibrium during physically demanding gameplay.

One might argue that even in traditional digital games, eliminating failure altogether would naturally elevate perceived competence and challenge–skill balance. However, the distinction in exergaming lies in the resolution of physiological and psychological asymmetry. While standard desktop games

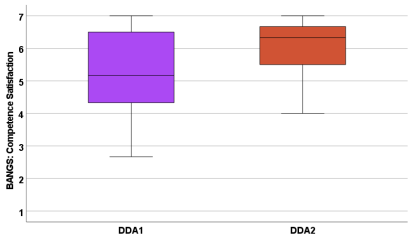


Fig. 3: Box plots for BANGS: Competence Satisfaction in DDA$_{Reactive}$ and DDA$_{Proactive}$ conditions.

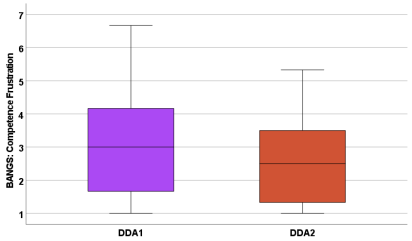


Fig. 4: Box plots for BANGS: Competence Frustration in DDA$_{Reactive}$ and DDA$_{Proactive}$ conditions.

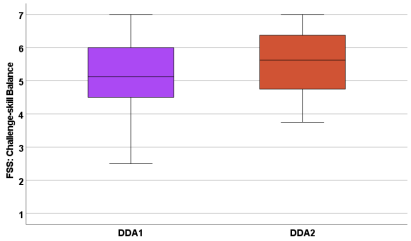


Fig. 5: Box plots for FSS: Challenge-skill Balance in DDA$_{Reactive}$ and DDA$_{Proactive}$ conditions.

TABLE II: Descriptive statistics and t-test results for resting and average heart rate in DDA$_{Reactive}$ and DDA$_{Proactive}$ conditions

Variables	**N**	**M ± SD**		**SEM**		**95% CI**		**t(df)**	***p***	***d***
		Reactive	Proactive	Reactive	Proactive	Reactive	Proactive			
Average Resting Heart Rate	36	84.57 ± 13.58	82.36 ± 12.88	2.26	2.15	[79.97, 89.16]	[78.00, 86.72]	0.766 (35)	0.449	0.13
Average Heart Rate	32	148.85 ± 16.43	146.13 ± 17.64	2.90	3.12	[142.92, 154.77]	[139.77, 152.49]	1.144 (31)	0.261	0.20

TABLE III: Internal consistency scores, descriptive statistics, and Wilcoxon signed-rank test results for variables under DDA$_{Reactive}$ and DDA$_{Proactive}$

Variables	**N**	**Cronbach's α**		**M ± SD**		**Mdn**		**IQR**		**Ties**	**Z**	**p**	**r**
		Reactive	Proactive	Reactive	Proactive	Reactive	Proactive	Reactive	Proactive				
BRPE: Perceived Exertion	36	–	–	14.42 ± 2.25	14.31 ± 2.34	14.00	14.00	2.75	3.00	10	-0.489	0.625	0.08
BANGS: Competence Satisfaction	36	0.88	0.79	5.28 ± 1.31	6.07 ± 0.81	5.17	6.33	2.26	1.25	6	-3.048	0.002*	0.51
BANGS: Competence Frustration	36	0.86	0.57	3.09 ± 1.61	2.51 ± 1.17	3.00	2.50	2.75	2.26	6	-2.052	0.040*	0.34
FSS: Challenge-Skill Balance	36	0.84	0.81	5.15 ± 1.00	5.57 ± 0.97	5.13	5.63	1.50	1.69	9	-2.388	0.017*	0.40

pause both cognitive processing and motor inputs upon failure, a DDA_{Reactive} penalty forces an abrupt cessation of in-game progress while the human body remains under being active. By implementing $DDA_{\text{Proactive}}$, we demonstrate the unique utility of anticipatory scaling in resolving this asymmetry. This approach allows players to maintain the flow of physical locomotion while simultaneously facilitating positive PX.

Our findings further suggest that the absence of repeated failure states in $DDA_{\text{Proactive}}$ does not necessarily reduce physical effort during gameplay [23] as no significant differences were observed in AHR or perceived exertion between conditions. Although missing heart rate data and the relatively small effect size for AHR warrant cautious interpretation, converging evidence from perceived exertion measures suggests that players in $DDA_{\text{Proactive}}$ were working just as hard as those in DDA_{Reactive}. These findings imply that repeated failure events may not be necessary to sustain physical engagement in exertion-based challenges.

Our results also indicate that repeated failures before DDA_{Reactive} activation increased competence frustration ($p = .040$, $r = .34$). One possible explanation is that primarily physical challenges may provide fewer opportunities for strategic adaptation than cognitively oriented gameplay, where failure can sometimes produce positively perceived frustration or reflective problem-solving [17, 41]. Unlike strategy-based game challenges, where players may reconsider tactics or develop alternative solutions after failure, exertion-based gameplay may rely more heavily on sustained physical output. Consequently, repeated punishments may contribute to heightened frustration [17, 18, 30, 31]. Although the BANGS frustration subscale demonstrated relatively low internal consistency, converging evidence from challenge–skill balance and competence satisfaction supports a consistent overall pattern favouring $DDA_{\text{Proactive}}$.

Taken together, these findings suggest several implications for exergame design. First, proactive adaptation may be particularly beneficial in exergames where physical exertion constitutes the primary challenge dimension. By reducing repeated failure events, pre-failure adaptation in $DDA_{\text{Proactive}}$ may help maintain competence satisfaction and challenge–skill balance without diminishing physical effort. Second, designers adopting reactive adaptation may benefit from incorporating additional challenge dimensions, such as cognitive tasks or strategic decision-making, alongside physical exertion. Providing alternative problem-solving pathways may allow players to recover from setbacks without relying exclusively on increased physical demand.

Importantly, these findings should not be interpreted as suggesting that $DDA_{\text{Proactive}}$ is universally superior to DDA_{Reactive}. Reactive adaptation may remain effective in exergames involving multiple challenge dimensions. However, in locomotion-based exergames where physical challenge constitutes the primary challenge dimension, relying exclusively on failure-driven difficulty adjustment may contribute to cycles of

frustration. In these contexts, $DDA_{\text{Proactive}}$ appears to provide a more supportive PX while preserving physical effort.

A. Limitations and Future Work

While this study provides empirical evidence regarding the effects of DDA functions in a running exergame, several limitations should be acknowledged. First, technical issues resulted in missing AHR data for four participants. Although perceived exertion provided converging evidence, future work should use more robust physiological logging methods. Second, despite using RHR to assess washout adequacy and order effects, residual fatigue between sessions may still have influenced responses. Future studies should consider alternative measures of physical effort, such as normalised or multimodal physiological metrics.

Third, the findings are situated within a primarily locomotion-based exergame context that isolated physical challenge to continuous running in place. Consequently, the relationship between setbacks, physical effort, and PX may differ in exergames involving more complex embodied interactions or movement mechanics. Similarly, the present study focused mainly on physical challenge, and it remains unclear how DDA would influence PX in hybrid exergames combining physical, cognitive, or emotional challenges [16].

Finally, this study adopted a single-session design with a moderate sample size. Longitudinal studies are needed to determine whether the observed benefits of $DDA_{Proactive}$ persist over time. Although $DDA_{Proactive}$ improved short-term competence satisfaction and challenge–skill balance, reducing failure states too aggressively may also reduce longer-term suspense or motivational tension.

IX. Conclusion

This study compared DDA_{Reactive} and $DDA_{\text{Proactive}}$ in a running exergame to examine their effects on PX and physical effort. The findings suggest that $DDA_{\text{Proactive}}$ improved competence satisfaction and challenge–skill balance while reducing competence frustration, without reducing physical effort. Furthermore, the study indicates that even relatively simple DDA implementations using real-time player data can effectively support challenge balance and positive PX. These findings are particularly relevant for exergame design, where sustaining engagement while preserving physical benefits is important.

References

[1] M. D. Finco and R. W. Maass, "The history of exergames: promotion of exercise through body interaction," in *Proc. IEEE SeGAH*, 2014, pp. 1–6.

[2] J. Marshall and C. Linehan, "Are exergames exercise? a scoping review of the short-term effects of exertion games," *IEEE Trans. Games*, vol. 13, no. 2, pp. 160–169, 2020.

[3] F. Mueller *et al.*, "Designing sports: a framework for exertion games," in *Proc. CHI*, 2011, pp. 2651–2660.

[4] E. A. Boyle, T. M. Connolly, T. Hainey, and J. M. Boyle, "Engagement in digital entertainment games: A systematic review," *Comput. Human Behaviour*, vol. 28, no. 3, pp. 771–780, 2012.

[5] J. Sinclair, P. Hingston, and M. Masek, "Considerations for the design of exergames," in *Proceedings of the 5th international conference on Computer graphics and interactive techniques in Australia and Southeast Asia*, 2007, pp. 289–295.

[6] ——, "Exergame development using the dual flow model," in *Proceedings of the Sixth Australasian Conference on Interactive Entertainment*, 2009, pp. 1–7.

[7] A. L. Martin-Niedecken and U. Götz, "Go with the dual flow: evaluating the psychophysiological adaptive fitness game environment "plunder planet"," in *Serious Games: Third Joint International Conference, JCSG 2017, Valencia, Spain, November 23-24, 2017, Proceedings 3*. Springer, 2017, pp. 32–43.

[8] M. Csikszentmihalyi, M. Csikszentmihalyi, S. Abuhamdeh, and J. Nakamura, "Flow," *Flow and the foundations of positive psychology: The collected works of Mihaly Csikszentmihalyi*, pp. 227–238, 2014.

[9] M. Zohaib, "Dynamic difficulty adjustment (dda) in computer games: A review," *Advances in Human-Computer Interaction*, vol. 2018, pp. 1–12, 11 2018.

[10] D. Dziedzic, "Dynamic difficulty adjustment systems for various game genres," *Homo Ludens*, vol. 9, no. 1, pp. 35–51, 2016.

[11] J. D. Smeddinck, R. L. Mandryk, M. V. Birk, K. M. Gerling, D. Barsilowski, and R. Malaka, "How to present game difficulty choices? exploring the impact on player experience," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2016, pp. 5595–5607.

[12] S. Chorrojprasert, "Exploring the effects of dynamic difficulty adjustment functions in running exergames on player experience and physical effort," Ph.D. dissertation, Univ. of York, 2025.

[13] A. L. Martin-Niedecken and E. D. Mekler, "The exercube: participatory design of an immersive fitness game environment," in *Joint international conference on serious games*. Springer, 2018, pp. 263–275.

[14] A. Darzi, S. M. McCrea, D. Novak *et al.*, "User experience with dynamic difficulty adjustment methods for an affective exergame: Comparative laboratory-based study," *JMIR Serious Games*, vol. 9, no. 2, p. e25771, 2021.

[15] P. Moschovitis and A. Denisova, "Keep calm and aim for the head: Biofeedback-controlled dynamic difficulty adjustment in a horror game," *IEEE Transactions on Games*, 2022.

[16] A. Denisova, C. Guckelsberger, and D. Zendle, "Challenge in digital games: Towards developing a measurement tool," in *Proceedings of the 2017 chi conference extended abstracts on human factors in computing systems*, 2017, pp. 2511–2519.

[17] J. w. Frommel, M. Klarkowski, and R. L. Mandryk, "The struggle is spiel: On failure and success in games," in *Proc. FDG*, 2021, pp. 1–12.

[18] W. Van den Hoogen, K. Poels, W. IJsselsteijn, and Y. De Kort, "Between challenge and defeat: Repeated player-death and game enjoyment," *Media Psychol.*, vol. 15, no. 4, pp. 443–459, 2012.

[19] S. Petralito *et al.*, "A good reason to die: how avatar death and high challenges enable positive experiences," in *Proc. CHI*, 2017, pp. 5087–5097.

[20] J. Juul, *The art of failure: An essay on the pain of playing video games*. MIT Press, 2013.

[21] R. M. Ryan, C. S. Rigby, and A. Przybylski, "The motivational pull of video games: A self-determination theory approach," *Motivation and emotion*, vol. 30, pp. 344–360, 2006.

[22] J. Juul, "In search of lost time: on game goals and failure costs," in *Proc. FDG*, 2010, pp. 86–91.

[23] ——, "Fear of failing? the many meanings of difficulty in video games," *Video Game Theory Reader 2*, pp. 237–252, 2009.

[24] M. A. M. Cuerdo, A. Mahajan, J. Mao, and E. F. Melcer, "Try again?: A macro-level taxonomy of the challenge and failure process in games," in *Proc. IEEE CoG*, 2023, pp. 1–8.

[25] B. Aytemiz and A. M. Smith, "A diagnostic taxonomy of failure in videogames," in *Proc. FDG*, 2020, pp. 1–11.

[26] C. G. Anderson, K. Campbell, and C. Steinkuehler, "Building persistence through failure: the role of challenge in video games," in *Proc. FDG*, 2019, pp. 1–6.

[27] K. Waldenmeier *et al.*, "Cheat codes as external support for players navigating fear of failure in digital games," in *Proc. CHI*, 2024, pp. 1–13.

[28] R. Furze, "Challenge," in *Routledge Companion to Video Game Studies*. Routledge, 2014, pp. 143–151.

[29] C. Foch and B. Kirman, ""slow down and look": Desirable aspects of failure in video games," in *Proc. FDG*, 2021, pp. 1–10.

[30] K. M. Gilleade and A. Dix, "Using frustration in the design of adaptive videogames," in *Proc. ACE*, 2004, pp. 228–232.

[31] F. Allison, M. Carter, and M. Gibbs, "Good frustrations: The paradoxical pleasure of fearing death in DayZ," in *Proc. OzCHI*, 2015, pp. 119–123.

[32] D. Haglind, "Coping with success and failure–a qualitative study on athletes and coaches in track and field," Technical Report, 2003.

[33] M. Goršič, A. Darzi, and D. Novak, "Comparison of two difficulty adaptation strategies for competitive arm rehabilitation exercises," in *2017 international conference on rehabilitation robotics (ICORR)*. IEEE, 2017, pp. 640–645.

[34] A. L. Martin-Niedecken, A. Mahrer, K. Rogers, E. D. de Bruin, and A. Schättin, ""hiit" the exercube: comparing the effectiveness of functional high-intensity interval training in conventional vs. exergame-based training," *Frontiers in Computer Science*, vol. 2, p. 33, 2020.

[35] A. Seyderhelm and K. Blackmore, "Systematic review of dynamic difficulty adaption for serious games: the importance of diverse approaches," *Available at SSRN*, vol. 3982971, 2021.

[36] D. Ang, "Difficulty in video games: understanding the effects of dynamic difficulty adjustment," in *Proc. C&C*, 2017, pp. 544–550.

[37] L. Climent, A. Longhi, A. Arbelaez, and M. Mancini, "A framework for designing reinforcement learning agents with dynamic difficulty adjustment in single-player action video games," *Entertainment Computing*, vol. 50, p. 100686, 2024.

[38] D. S. Lora-Ariza, A. A. Sánchez-Ruiz, P. A. González-Calero, and I. Camps-Ortueta, "Measuring control to dynamically induce flow in Tetris," *IEEE Trans. Games*, vol. 14, no. 4, pp. 579–588, 2022.

[39] O. S. Ajani and R. Mallipeddi, "Pareto-based dynamic difficulty adjustment of a competitive exergame for arm rehabilitation," *Int. J. Hum.-Comput. Stud.*, vol. 178, p. 103100, 2023.

[40] A. L. Martin-Niedecken, K. Rogers, L. Turmo Vidal, E. D. Mekler, and E. Márquez Segura, "Exercube vs. personal trainer: evaluating a holistic, immersive, and adaptive fitness game setup," in *Proceedings of the 2019 CHI conference on human factors in computing systems*, 2019, pp. 1–15.

[41] M. Wolterink and S. Bakkes, "Towards explainable prediction of player frustration in video games," in *Proc. FDG*, 2021, pp. 1–10.

[42] M. Jetté, K. Sidney, and G. Blümchen, "Metabolic equivalents (METs) in exercise testing, exercise prescription, and evaluation of functional capacity," *Clin. Cardiol.*, vol. 13, no. 8, pp. 555–565, 1990.

[43] Q. Mi and T. Gao, "Adaptive rubber-banding system of dynamic difficulty adjustment in racing games," *ICGA J.*, vol. 44, no. 1, pp. 18–38, 2022.

[44] M. Ketcheson, L. Walker, and T. N. Graham, "Thighrim and calf-life: a study of the conversion of off-the-shelf video games into exergames," in *Proceedings of the 2016 chi conference on human factors in computing systems*, 2016, pp. 2681–2692.

[45] G. Borg, *Borg's perceived exertion and pain scales.* Human kinetics, 1998.

[46] N. Ballou, A. Denisova, R. Ryan, C. S. Rigby, and S. Deterding, "The basic needs in games scale (bangs): A new tool for investigating positive and negative video game experiences," *International Journal of Human-Computer Studies*, vol. 188, p. 103289, 2024.

[47] S. A. Jackson and H. W. Marsh, "Development and validation of a scale to measure optimal experience: The flow state scale," *Journal of sport and exercise psychology*, vol. 18, no. 1, pp. 17–35, 1996.