Investigating the Relationship Between the Scale of Reflex-Based Challenge and Flow, Immersion, and Game Experience

Zainab Ezzeddine,
School of Computing, Engineering and the Built Environment
Edinburgh Napier University, Edinburgh
UK
zainab.ezz@gmail.com

John Paul Vargheese,
School of Computing, Engineering and the Built Environment
Edinburgh Napier University, Edinburgh,
UK
jpvargheese@acm.org

Emilia Sobolewska,
School of Computing, Engineering and the Built Environment
Edinburgh Napier University, Edinburgh,
UK
e.sobolewska@napier.ac.uk

This study examines how varying levels of reflex-based challenge influence player experience in *Geometry Dash*, a non-narrative rhythm game. Using a within-subjects design with *N* = 18 participants (predominantly aged 26 – 33, beginner to intermediate gaming ability), we measured Flow, sensory immersion, and affective states across easy, medium, and difficult game levels using the Game Experience Questionnaire (GEQ). Results indicated that Flow peaked at medium difficulty consistent with Csikszentmihalyi's optimal challenge hypothesis. However, a repeated-measures ANOVA test indicated no significant difference between the medium and difficult conditions. Sensory Immersion remained stable across conditions, while Tension/Annoyance increased linearly with difficulty. Multiple Regression analyses showed challenge consistently predicted Flow, whereas Competence only became significant in difficult conditions. These findings suggest that reflex-based tactical challenge follows distinct patterns that may influence player experience in mechanically focussed, non-narrative games, with implications for designing skill-based games that seek to balance challenge and enjoyment. Methodologically, sensory immersion was analysed separately from imaginative immersion to better reflect the reflex-based gameplay.

Immersion. Flow. Reflex-based tactical challenge. Optimal experience.

1. INTRODUCTION

In game studies, game immersion is widely conceptualised as a psychological state in which players are deeply engaged with a game's world (Michailidis, Balaguer-Ballester, & He, 2018). Further, immersion encompasses a spectrum of cognitive and emotional involvement, particularly in relation to gameplay mechanics and challenges (McMahan, 2003). Although immersion has been linked to narrative (Lu & Moller, 2024), graphical fidelity (Gerling et al., 2013), and game mechanics (Argasiński & Węgrzyn, 2018), a crucial – yet underexplored – aspect is the role of specific types of tactical challenges. While research is growing (e.g., Kuo et al., 2024), there remains a gap in isolating reflex mechanics from narrative. In this study, reflex-based tactical challenge is characterised by the levels of speed, and the frequency of time-precision obstacles requiring immediate single-tap motor responses, and serving as the operationalised independent variables. This taxonomy builds on the framework of Denisova et al. (2020), which conceptualises challenge as a multi-dimensional construct, explicitly separating rule-based cognitive demands from immediate reflex and motor-skill requirements. In Denisova et al.'s study (2020), this mechanical demand is understood as a performative challenge. Blackman (2024) highlights that while presence and the spatiality of being there are central to virtual worlds, immersion in non-narrative contexts often relies on the performative nature of the gameplay itself. In this paper, Denisova et al.'s (2020) definition of motor challenge outlines the quantitative structural constraints of the game level, while Blackman's (2024) performance lens captures the active, tactical choices through which a player exercises agency to navigate those constraints and sustain presence. Adams (2004) states that tactical challenges which foster immersion are often simple, but require precision, game control and consistency without interruptions. Therefore, challenge and immersion remain fundamental components of player engagement (Umar et al., 2022). Another way immersion is defined in video games is through three psychological states of deep focus: engagement, engrossment, and total

© Ezzeddine et al. Published by BCS Learning and Development Ltd. Proceedings of BCS HCI 2026, UK

immersion (Brown & Cairns, 2004). Jennett et al. (2008) claim that Immersion and Flow overlap in losing track of time and engaging challenges, but, that immersion is evidently a precursor for Flow. Recent research by Kuo et al (2024) supports this by demonstrating that the mastery of game challenges serves as a primary predictor for both engrossment and total immersion. This framework is particularly valuable for studying a reflex-based challenge game like *Geometry Dash*, which lacks virtual presence but relies on mechanical absorption. While these frameworks describe immersion broadly, less is known about how reflex-based tactical challenges, specifically precision mechanics, drive immersion and Flow in non-narrative games. To address this, the present study engages with Csikszentmihalyi's (1990) Flow theory and optimal experience concept, the SCI-model of immersion (Ermi & Mäyrä, 2005), and Adams' (2004) types of immersion. Specifically, the 2D mobile game *Geometry Dash* is used to examine these dynamics, as its rhythm-based mechanics and customisation nature provide a pure environment for reflex-based challenge. By focusing on tactical demands involving speed, time-precision, and obstacles, this study examines the role of mechanical intensity in shaping player immersion, Flow, and game experience (Umar et al., 2022).

2. LITERATURE REVIEW

2.1 Challenge-based Immersion

Immersion in digital gaming is traditionally categorised into experiential and technological perspectives (Cairns et al., 2014). While the technological perspective emphasises virtual presence through high-fidelity graphics, Blackman (2024) highlights that immersion in non-narrative contexts relies on the performative nature of the gameplay itself rather than visual presence. To isolate these mechanics, this study utilises Ermi and Mäyrä's (2005) three-dimensional SCI-model, which identifies sensory, imaginative, and challenge-based immersion as distinct pillars. Umar et al. (2022) have recently validated this model as a robust framework for identifying immersive features in modern gaming environments. In the context of this study, the focus lies predominantly in challenge-based immersion. By utilising a 2-dimensional visual design and cancelling all music and sound effects, the imaginative and auditory sensory elements are suppressed. This methodological choice is supported by Umar et al. (2022), whose research indicates that immersion is a multi-dimensional state where individual features can be isolated to empirically determine their specific impact on player engagement. This aligns with Blackman's (2024) assertion that the performative spatiality of a player's actions, the rhythm of movement and reaction, can sustain immersion even in the absence of high-fidelity environmental cues. This further connects to Adams' (2004) framework of tactical immersion, which occurs when players are fully absorbed in gameplay requiring quick decision-making and motor skills.

2.2 Measuring Reflex-Based Challenge and Flow

Csikszentmihalyi (1990) defines Flow as an optimal experience that leaves players "so involved in the activity that nothing else seems to matter". Formally, this is a phenomenological state of autotelic engagement, where the player's consciousness is occupied by a narrow field of stimuli (Procci & Bowers, 2011). Following Chen (2007), Flow arises when challenge and ability are balanced. McDaniel, Davis, and Weeks (2024) reinforce this by demonstrating through simulation that maintaining the 'Flow channel' requires the game to properly increase difficulty to keep pace with player growth. Without this calibration, players slip into frustration (when difficulty outweighs skill) or boredom (when skill outweighs difficulty). While critics claim that Flow is an 'all-or-nothing state' (Cairns, Cox, & Nordin, 2014), modern research confirms that Flow dimensions are not absolute and can fluctuate (Michailidis, Balaguer-Ballester, & He, 2018). This study builds on Kuo et al. (2024), whose research indicates that mastery of tactical challenges, specifically successful game completion, strongly drives engrossment and total immersion. In reflex-based games like *Geometry Dash*, this implies that immersion is a dynamic result of the player's ability to execute precise motor responses under time-pressure. This is further supported by Lora-Ariza et al. (2022), who suggest that a player's ability to maintain behavioural control over mechanical demands is a primary indicator of entering the Flow state. Furthermore, this research builds on neurocognitive work by Klasen et al. (2012). While neuroimaging requires controlled gameplay and post-scan self-reports due to fMRI scanner constraints, our design complements these insights through active, real time flow measurements. Additionally, while neuroimaging protocols rely on broader difficulty tiers to maintain baseline stability, our approach uses discrete, trackable reflex parameters (speed and obstacle density) to aim for a more ecologically valid mapping of the Flow state.

2.3 Real-time Experience and Competence

To address limitations in state-specific skill measurement found in prior work (Procci & Bowers, 2011; Rutrecht et al., 2021), this research operationalises discrete tactical challenges. While Rutrecht et al. (2021) found higher levels of Flow in rhythm-based action games are associated with a subjective acceleration of time, their reliance on post-play surveys conflicts with Experience Sampling Method (ESM) principles (Csikszentmihalyi and Larson, 2014). As noted by

Csikszentmihalyi and Larson (2014), capturing in-the-moment experience is critical to avoiding the bias of retrospective reporting. By using objectively set speeds in *Geometry Dash* and administering the Game Experience Questionnaire (GEQ) immediately after each level, this study makes the effects of reflex-based tactical challenge trackable. This methodological decision to track performance parallels the objective control metrics proposed by Lora-Ariza et al. (2022) to evaluate the Flow channel. This approach is further refined by the distinction between avatar skill (character upgrades) and player skill (competence), a concept expanded upon by McDaniel, Davis, and Weeks (2024). They argue that player skill involves intricate factors such as reaction time and specific game knowledge, whereas avatar skill encompasses statistical upgrades. By controlling for avatar upgrades, this study seeks to isolate player skill as the primary variable. This builds on Cox et al.'s (2012) findings, which established that time pressure significantly increases immersion, while ensuring the moderation effect of expertise is measured through state-specific competence rather than general gaming habits.

2.4 Theoretical Synthesis

This study employs a controlled experiment to isolate tactical immersion (Adams, 2004) through reflex-based challenges in a non-narrative, visually minimal 2D game. By limiting sensory variables (muting audio and choosing neutral, consistent colours), the design focuses purely on challenge-based immersion (Ermi & Mäyrä, 2005). Integrating foundational Flow theory with the SCI-model (Ermi and Mäyrä, 2005; Umar et al., 2022) and the mastery of tactical challenges (Kuo et al., 2024), this study extends previous work that adopted a purely correlational approach. By relating GEQ data with controlled difficulty parameters, this research aims to examine how reflex-based challenges influence Flow, Immersion, and player experience. This is intended to bridge the gap between the performative gameplay described by Blackman (2024) and the foundational optimal experience, providing a precise mapping of how visual-only, time-precision challenges influence the player experience.

2.5 Research Questions and Hypotheses

Based on the performative boundaries and optimal experience frameworks reviewed above, this paper addresses the following research questions and evaluates their corresponding hypotheses:
RQ1: How do different levels of reflex-based tactical challenge affect players' Flow?
H1: Players will report greater Flow in the medium-difficulty condition compared to easy or difficult, measured by the GEQ Flow subscale.
RQ2: How do different levels of reflex-based tactical challenge affect players' sensory immersion?
H2: Players will report higher sensory immersion in the medium-difficulty condition than in easy or difficult conditions, measured by the GEQ Sensory Only subscale.
RQ3: How do different levels of reflex-based tactical challenge affect players' game experience (tension, positive affect, and negative affect)?
H3: Players will experience the most positive feelings in the medium-difficulty condition, but more negative emotions, tension, and annoyance in the difficult level.

3. METHODOLOGY

3.1 Design

This study used a within-subject experimental design procedure using *Geometry Dash*, a 2D mobile-game, alongside the Game Experience Questionnaire (GEQ) to deduce the effect of levels of difficulty (easy, medium, difficult) in terms of speed, time-precision, and obstacles, on Flow, sensory/imaginary immersion, challenge, competence, tension/annoyance, and positive and negative affect experienced by participants. By exposing each participant to all conditions, the within-subject design supports causal inference while controlling for individual differences in baseline player skill. These difficulty conditions were selected to provide progressive increasing levels of reflex-based challenge while remaining consistent with previous Flow research comparing, low, moderate and high challenge conditions.

3.2 Participants

Through opportunistic sampling, 18 participants aged between 19 and 47 years (M = 29.5, SD = 6.4) were recruited from a university campus. Although participants were recruited from a single university campus, the sample included participants from a range of national and gaming backgrounds (see Table 1). Although the sample consists of 18 participants, the within-subject design increases statistical power by using each participant as their own control, which minimises the impact of individual differences in baseline gaming skill. Following recommended sample size outlined by Lakens (2022), which builds directly on the classic within-subject power models established by Maxwell, Delaney and Kelley (2017), this study design produces 54 experimental sessions across the three difficulty conditions. As shown in Table 1, the cohort represents a broad range of experience levels, with the majority identifying as beginners, novices, or intermediates. This ability metric was captured via a pre-study questionnaire where participants self-reported their general expertise with playing digital games, rather than track-specific proficiency within *Geometry Dash* or the rhythm

platformer genre.

Table 1: *Participant demographics and expertise*

Variable	Category	*n*	%
Gender	*Male*	*8*	*44.4*
	Female	*9*	*50.0*
	Gender Neutral	*1*	*5.6*
Nationality	*Scottish*	*7*	*38.9*
	Nigerian	*4*	*22.2*
	Ukrainian	*2*	*11.1*
	Greek	*1*	*5.6*
	Indian	*1*	*5.6*
	Italian	*1*	*5.6*
	Lebanese	*1*	*5.6*
	Thai	*1*	*5.6*
Game Ability	*Novice*	*4*	*22.2*
	Beginner	*5*	*27.8*
	Intermediate	*7*	*38.9*
	Advanced	*1*	*5.6*
	Expert	*1*	*5.6*
Age	*Means (SD)*	*29.5 (6.4)*	-
	Range	*19 - 47*	-

3.3 Materials and Apparatus

3.3.1. The video game platform – Geometry Dash
As a non-narrative, 2D, rhythm-based, and side-scrolling platformer game, *Geometry Dash* is dominated by sensory and challenge-based immersion (Ermi & Mäyrä, 2005). *Geometry Dash* is a fast-paced, reflexive game which incorporates time-precision based challenge by having only a single-tap control scheme (1 tap = 1 jump control function) (RobTop Games, 2013) (see Figure 1). Jumping over obstacles is controlled by timing single-tab jumps accurately; relying on both visual and auditory cues to avoid hazards and reach the endpoint (Li & Raferty). Furthermore, the single-tap control was kept as is in the settings to eliminate complex input and retain friendly learnability for people from all game ability levels. As Adams (2004) describes, tactical immersion occurs when players enter an automatic state of perfect timing and reaction, which is known as being 'in the zone'. Crucially, for this study's purpose, *Geometry Dash* was chosen because it provides the option to create customised levels with different speed levels and number of obstacles/jumps (see Figures 2 and 3). This allows for the creation of three levels of difficulties, easy, medium, and difficult, allowing for a progressive challenge curve (Ermi & Mäyrä, 2005; RobTop Games, 2013). Moreover, with the availability of 'creating' your own levels on *Geometry Dash,* audio (Haehn, Böffel, & Schlittmeier, 2023; Guillen, Jylhä, & Hassan, 2021) and colours (Canossa, Badler, Magy, & Anderson, 2016) can affect immersion, Flow, and player's affective game experience further. This can also be controlled for by muting the songs and sound effects (see Figure 4) and choosing different shades of white, black, and grey for background colours (see Figure 3).

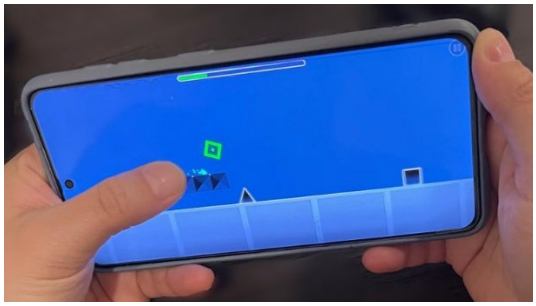


Figure 1: *Geometry Dash gameplay showing the single-tap jump mechanic*



Figure 2: *Creating level on Geometry Dash screenshot*



Figure 3: *Geometry Dash level settings*

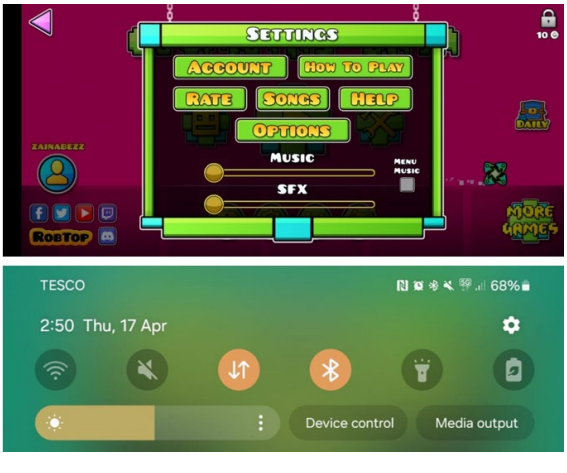


Figure 4: *Controlling for audio distractions*

3.3.2. Game Experience Questionnaire (GEQ)
In 2007, IJsselsteijn, de Kort, Poels, and Jurgelionis developed the GEQ to provide a comprehensive and reliable assessment of digital game experiences which is also robust and applicable, drawing mainly from the SCI-model by Ermi and Mäyrä (2005) and the *GameFlow* model by Sweetser and Wyeth (2005). Alongside these established theoretical models, large-scale field trials with both casual and hardcore gamers were applied to inform the structure and content of the GEQ. The questionnaire encompasses a range of experiential dimensions, ensuring both subjective and reflexive responses to gameplay are systematically evaluated, both during and after gameplay (IJsselsteijn, de Kort, Poels, & Jurgelionis, 2007). Hence, the GEQ comprises three sequential modules: the core questionnaire (measuring immersion, Flow, competence, positive and negative affect, tension/annoyance, and challenge), the social presence module, and the post-game module, all using five Likert-scale items per component (plus a spare for translation reliability). In 2018, Johnson, Gardner, and Perry validated the GEQ and claimed that the Flow, immersion, competence, and positive affect constructs are empirically supported, whereas the rest, negative affect, tension/annoyance, and challenge (which they advise should be combined into a single negativity construct), are not, with signs that they might be conceptually overlapping with other constructs.

3.4 Procedure
3.4.1. Environmental controls
The experiment was conducted in a controlled, isolated setting to minimise environmental variables (see Figure 5). To isolate challenge-based immersion, the smartphone was muted and all notifications were disabled (see Figure 4). This methodological choice follows Umar et al. (2022), ensuring that immersion is driven by performative spatiality rather than auditory cues.



Figure 5: *Controlled setting*

3.4.2. Pilot study and levels' iterations
The levels were designed using *Geometry Dash's* customisation features to isolate reflex-based difficulty through speed and obstacle density. A pilot evaluation was conducted with the first five participants to assess the experimental setup and validate the progression of the 'Flow channel' as defined by McDaniel et al. (2024). Based on participant feedback, the medium and difficult levels were adjusted to establish a clearer progression in tactical challenge. While these initial five participants completed the pre-calibration track iterations and did not repeat the study following refinements, their data was retained in the complete dataset to serve as a baseline comparison. The remaining 13 participants completed the main study using the revised tracks. Level order was counterbalanced across all participants to minimise order effects (Charness, Gneezy, & Kuhn, 2012), while the within-subject design controlled for individual baseline differences. By using the default *Geometry Dash* speeds (1×, 2×, and 3×) and varying the jump frequency, the researchers created discrete 'tactical tiers' that map these levels onto Denisova et al.'s (2020) framework of motor challenge and scaling the overall difficulty. As shown in Table 2, the Medium level, which applies speed 2×, had the number of required jumps reduced from 21 to 18 jumps, and the Difficult level, which applies speed 3×, increased the number of required jumps from 17 to 22 (To see the changes in the levels' layouts, see Appendix 1). These adjustments were necessary to prevent participants from exiting the 'Flow channel' due to perceived unfairness in the medium condition.

Table 2: *Tactical challenge parameters by level*

Difficulty	Speed Level	Initial Jumps	Final Jumps (Post-pilot)
Easy	*1×*	*17*	*17*
Medium	*2×*	*21*	*18*
Difficult	*3×*	*17*	*22*

3.4.3. Experimental Procedure
Participants were briefed on the study's objectives, confidentiality, and safety risks associated with rapid visual stimuli. After providing consent, the researcher demonstrated the single-tap mechanic. To ensure adequate immersion without inducing fatigue, each level required a minimum of 2 minutes and a maximum of 8 minutes of play. Following a within-subject design, participants completed the GEQ-Core questionnaire after each level, with level order shuffled to minimise order bias (Charness, Gneezy, & Kuhn, 2012). The three customised game levels (controlling for in game sound effects, music, and background colours) and the experimental execution (a controlled setting alongside counterbalancing level order) were designed to minimise external noise and eliminate in game confounding variables

4. RESULTS AND ANALYSIS

This section presents descriptive statistics followed by repeated measures ANOVA to examine the effects of difficulty on Flow, Immersion, and Game Experience across the experimental conditions to address RQ1–RQ3. Additional analyses for Imaginary Immersion, Negative Affect, and Combined Immersion are reported in Appendix 2. Multiple regression analyses examine whether perceived Challenge and Competence are associated with Flow within each difficulty condition. Together, these analyses internally validate condition differences before evaluating predictive relationships among the study variables.

4.1 Descriptive Statistics

4.1.1. GEQ by difficulty level
As presented in Table 3, Imaginary Immersion, Sensory Immersion, Combined Immersion, and Flow all peaked at medium difficulty. Challenge increased with difficulty, while Competence declined as difficulty rose. Negative Affect was lowest at medium and highest at difficult levels, whereas Positive Affect was greatest at easy and lowest at difficult. Conversely, Tension/Annoyance was minimal at easy and maximal at difficult.

Table 3: GEQ subscale statistics by difficulty
Note. Values reported as Mean ± SD (0-4).

GEQ subscales	Difficulty levels		
	Easy	*Medium*	*Difficult*
Positive Affect	*2.8 ± 1.0*	*2.7 ± 1.1*	*2.3 ± 1.1*
Negative Affect	*0.8 ± 0.8*	*0.7 ± 0.7*	*0.9 ± 0.8*
Competence	*2.2 ± 1.3*	*1.7 ± 1.0*	*1.4 ± 1.2*
Sensory Immersion	*1.9 ± 1.0*	*2.2 ± 1.2*	*2.0 ± 1.1*
Imaginary Immersion	*1.2 ± 1.1*	*1.3 ± 1.0*	*0.9 ± 0.9*
Combined Immersion	*1.6 ± 1.0*	*1.7 ± 1.0*	*1.4 ± 0.9*
Flow State	*2.1 ± 1.0*	*2.6 ± 0.9*	*2.4 ± 0.8*
Challenge	*1.6 ± 1.1*	*2.3 ± 0.9*	*2.5 ± 1.0*
Tension/Annoyance	*0.8 ± 1.0*	*1.4 ± 1.1*	*1.7 ± 1.2*

4.2 Inferential Statistics

4.2.1. Flow
Difficulty level has a significant effect on Flow state, $F(2,34) = 4.38$, $p = .020$, $\eta^2 = .048$. Bonferroni-corrected pairwise tests showed significantly higher flow in Medium difficulty ($M_{Medium} = 2.6$, $SD = 0.9$) compared to Easy ($M_{Easy} = 2.1$, $SD = 1.0$; $p = .035$), but no differences between other conditions (all ps > .05) (For full repeated measures ANOVA models on secondary subscales, see Appendix 2).

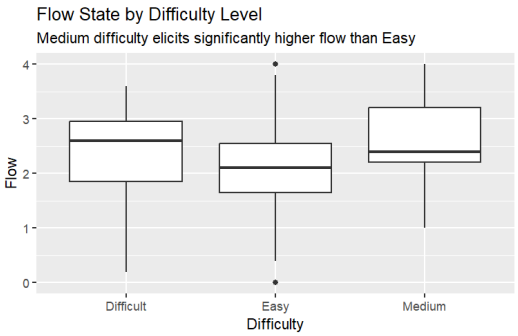


Figure 6: *Flow by difficulty level*

4.2.2. Challenge
Difficulty level had a significant and large effect on perceived Challenge, $F(2,34) = 13.75$, $p < .001$, $\eta^2 = .148$. Bonferroni-corrected comparisons showed players rated Easy difficulty ($M_{Easy} = 1.6$, $SD = 1.1$) as significantly less challenging than both Medium ($M_{Medium} = 2.3$, $SD = 0.9$; $p = .003$) and Difficult ($M_{Difficult} = 2.5$, $SD = 1.0$; $p = .001$) conditions. Conversely, there was no significant difference in perceived Challenge between the Medium and Difficult levels ($p = 1.000$).

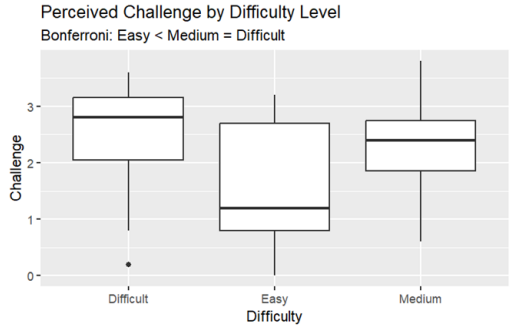


Figure 7: *Perceived Challenge by difficulty level*

4.2.3. Competence
Difficulty level had a significant effect on perceived Competence, $F(2,34) = 4.14$, $p = .024$, $\eta^2 = .060$. Although Competence was lower in the Difficult condition than the Easy condition, this pairwise comparison was not significant after Bonferroni correction, ($M_{Easy} = 2.2$, $SD = 1.3$) and Difficult ($M_{Difficult} = 1.4$, $SD = 1.2$; $p = .068$) conditions. Perceived Competence in the Medium difficulty ($M_{Medium} = 1.7$, $SD = 1.0$) did not significantly differ from either the Easy or Difficult conditions (both ps >.26).

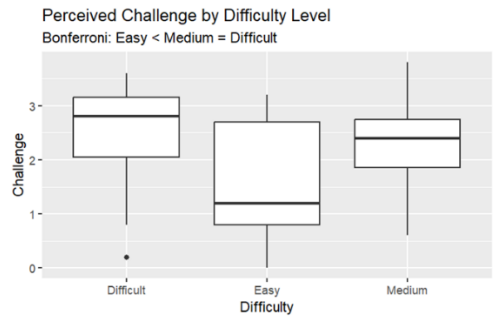


Figure 8: *Perceived Competence by difficulty level*

4.2.4. Sensory immersion

Difficulty level had no significant effect on Sensory Immersion**,** $F(1.34, 22.81) = 0.80$, $p = .416$, $\eta^2 = .011$. Bonferroni-corrected pairwise confirmed that there were no significant differences between Easy ($M_{Easy} = 1.9$, $SD = 1.0$), Medium ($M_{Medium} = 2.2$, $SD = 1.2$), and Difficult ($M_{Difficult} = 2.0$, $SD = 1.1$) conditions (all ps > .05). This finding should be interpreted with caution, however, due to marginal non-normality in the Medium condition.

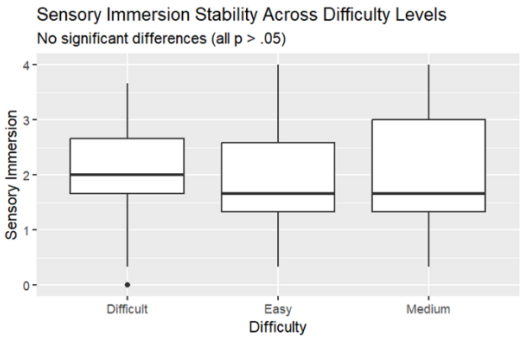


Figure 9: *Sensory immersion by difficulty levels*

4.2.5. Tension/Annoyance

Difficulty level had a significant effect on Tension/Annoyance, $F(2,34) = 6.22$, $p = .005$, $\eta^2 = .119$. Pairwise comparisons revealed significantly higher tension in the Difficult ($M_{Difficult} = 1.7$, $SD = 1.2$) condition compared to the Easy ($M_{Easy} = 0.8$, $SD = 1.4$; $p = .033$) condition. No significant differences were found between Difficult and Medium ($p = .296$) or Easy and Medium ($p = .102$) conditions. Due to deviations from normality in the Easy condition, these parametric results should be interpreted with caution.

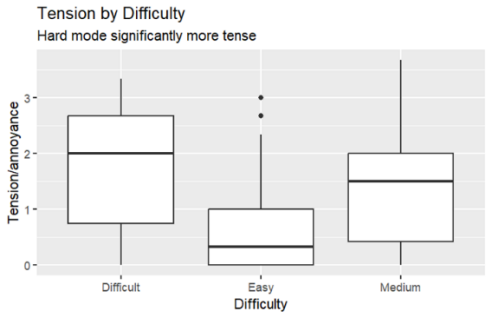


Figure 10: *Tension/Annoyance by difficulty level*

4.2.6. Positive affect

A Repeated-measures ANOVA with Greenhouse-Geisser correction ($\varepsilon = 0.739$) revealed a small but significant effect of difficulty on positive affect, $F(1.48, 25.13) = 3.91$, $p = .045$, $\eta^2 = .047$, after Mauchly's test indicated sphericity violation ($W = 0.647$, $p = .031$). While pairwise comparisons did not reach significance, a marginal decline emerged between Easy ($M_{Easy} = 2.8$, $SD = 1.0$) and Difficult ($M_{Difficult} = 2.3$, $SD = 1.1$; $p = .092$) conditions. Medium difficulty ($M_{Medium} = 2.7$, $SD = 1.1$) did not significantly differ from other conditions (both ps > .31). The Easy condition showed marginal non-normality ($p = .039$), and model residuals deviated from normality ($W = 0.945$, $p = .015$), suggesting cautious interpretation.

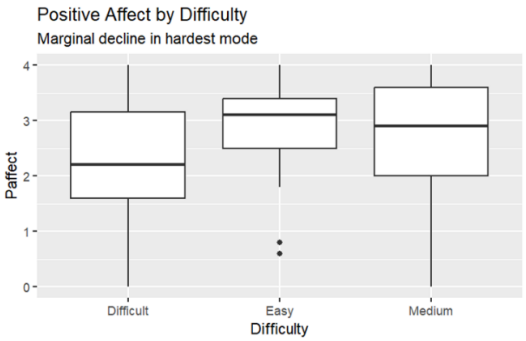


Figure 11: *Positive affect by difficulty level*

In summary, difficulty significantly affected Flow, Challenge, Competence, Tension/Annoyance, and Positive Affect, whereas Sensory Immersion remained stable across all conditions. These findings address RQ1-RQ3 by demonstrating how difficulty influences player experience. The following regression analyses examine whether perceived Challenge and Competence were associated with Flow within each difficulty condition.

4.3 Multiple Regression Analysis

4.3.1. Regression analysis: predictors of Flow

Prior to conducting the regression analyses, assumption checks were performed for all difficulty levels. Assumption checks indicated no substantial violations of normality, homoscedasticity, multicollinearity, or predictor independence across the models, with variance inflation factors (VIFs) remaining below 1.5 (see Appendix 3)**.** To examine the predictors of Flow across different challenge contexts, multiple regression analyses were conducted for each difficulty level. The results, summarised in Table 4, indicate that the relative contribution of Challenge and Competence differed across the three difficulty conditions.

4.3.2. Summary of regression patterns

This section summarises the key patterns emerging from the results of the regression analysis:

Consistency of Challenge: Challenge served as a significant predictor of Flow across all three difficulty levels, maintaining a strong predictive role regardless of the game's complexity.

Conditional Role of Competence: While competence significantly predicted Flow in the Easy and Difficult conditions, it was not a significant predictor in the Medium condition. This suggests that the association and Flow may be more pronounced at the extremes of the difficulty spectrum.

Model Variance: The predictors explained the highest proportion of variance in the Easy (63%) and Difficult (65%) conditions, while the lower proportion explained in the Medium condition (32.3%) suggests other unmeasured factors may be associated with Flow at that level of engagement. Together, these findings indicate that Challenge was a consistent predictor of Flow in *Geometry Dash*, while the role of perceived Competence appeared particularly important in the high-difficulty condition.

5. DISCUSSION

This study investigated how varying levels of reflex-based tactical challenge affect Flow, Immersion, and player experience. While the results suggest that challenge was consistently associated with Flow, in non-narrative environments, they also reveal a complex relationship between difficulty, perceived competence, and emotional affect. The following sections interpret these findings against existing literature to address the study's core research.

5.1 Track Level Iteration and Flow Balance Check

This section evaluates whether our customised game levels sufficiently isolated reflex-based challenge to examine Csikszentmihalyi's Flow framework. Rather than evaluating software development choices, these operational controls isolated core gameplay mechanics. Using multiple regressions alongside repeated measures ANOVAs provides an internal validation of our dataset, confirming that Flow was directly generated by balancing player competence against track challenge to map successfully onto established psychological theories. Geometry Dash was selected to control for narrative effects, allowing immersion to be interpreted mainly as tactical rather than story-driven. Imaginary Immersion remained consistently low across all conditions (Table 3), suggesting that the reported immersion was driven largely by gameplay mechanics rather than virtual presence. This aligns with Blackman (2024), who argues that immersion in non-narrative environments relies on the performative nature of the gameplay rather than the spatiality of 'being there'. The regression analysis supports the game's difficulty structure by identifying challenge as a consistent predictor of Flow across all three difficulty conditions. Notably, competence was not a significant predictor of Flow at the medium difficulty level. This finding suggests that at the optimal point of engagement, players may experience what Csikszentmihalyi (1990) describes as a 'loss of self-consciousness'. In this state, the player is so focused on the immediate requirements of rhythm and speed that they stop consciously monitoring their own skill. This aligns with the validation of the SCI-model by Umar et al. (2022), which identifies challenge-based immersion as a distinct pillar that can be isolated from sensory or imaginative elements. When challenge and perceived Competence are suitably balanced, the 'self' may recede, as the player simply reacts to the game's obstacles. This implies that at peak Flow, the external demand of the challenge handles the optimal experience automatically, making internal evaluations of competence secondary to the action itself.

Table 4: *Regression coefficients for Flow*

Condition	Predictor	β	SE	t	p	Adj. R^2
Easy	*Challenge*	*.862***	*.155*	*5.56*	*< .001*	*.630*
	Competence	*.323**	*.135*	*2.40*	*.030*	
Medium	*Challenge*	*.728***	*.229*	*3.18*	*.006*	*.323*
	Competence	*.313*	*.195*	*1.60*	*.130*	
Difficult	*Challenge*	*.658****	*.131*	*5.04*	*< .001*	*.650*
	Competence	*.450****	*.101*	*4.45*	*< .001*	

5.2 RQ1 – Flow and Difficulty

In line with Hypothesis 1, Flow reached its highest level in the medium-difficulty condition. While the difficult condition also produced high Flow, the absence of a significant difference between the two is consistent with a potential a 'Flow plateau' in reflex-based games. This may indicate that once a certain challenge threshold is met, Flow remains relatively stable even as difficulty pushes toward the player's upper limits. This aligns with Csikszentmihalyi's (1990) theory, where the 'Flow channel' represents a zone of engagement rather than a single point. McDaniel et al. (2024) reinforce this by demonstrating that maintaining this channel requires difficulty to keep pace with player growth. Both medium and difficult levels appear to provide the necessary tactical demand, defined by Adams (2004) as precision, timing, and consistency, to maintain deep engagement. Alternatively, the lack of variance may indicate that the sample size (N = 18) was underpowered to detect subtle statistical differences between the medium and difficult tiers. As established by Cox, Cairns, Shah and Carroll (2012), immersion in action games is heavily driven by intense time pressure and mechanical demand rather than visual context. Because *Geometry Dash* relies entirely on basic jump and timing mechanics to isolate this performative challenge (Denisova, Cairns, Guckelsberger and Zendle, 2020), player experiences across the two fastest speeds look highly similar. Detecting subtle variations within such closely aligned mechanical conditions may require a larger participant sample. Therefore, while our

sample of 18 participants has enough statistical power to show large primary trends within-subject, it remains bounded when separating the close experiential differences found at the upper limits of reflex-based challenge.

5.3 RQ2 – Sensory Immersion and Difficulty

Hypothesis 2 which posited that Flow would drive higher sensory immersion, was not supported. Sensory Immersion (isolated to visual stimuli) remained stable, confirming the performative nature of the task (Blackman, 2024). Although the highest level of sensory immersion was reported in the medium condition, the variation across levels was negligible. This suggests that in mechanically-driven games, sensory immersion is likely more linked to the constant audiovisual qualities of the game, such as the 2D side-scrolling and rhythmic stimuli, than to fluctuations in Flow. This supports the assertion by Kuo et al. (2024) that while mastery of game challenges drives engrossment, certain immersive states can be decoupled from the challenge-skill balance. These results align with findings by Jennett et al. (2008) and Ermi and Mäyrä (2005), suggesting that sensory absorption can be decoupled from the challenge-skill balance in high-speed, reflex-dependent environments. Furthermore, as described by Brown and Cairns (2004), while players reached engrossment, the total immersion state remained elusive due to the lack of narrative presence.

5.4 RQ3 – Affective Game Experience and Difficulty

The analysis of player affect yielded a nuanced picture of the optimal experience. Positive affect was highest in the easy condition, suggesting an 'enjoyability threshold' (Chen, 2007) where low demand allows for relaxation. However, this positive feeling decreased as the game became more demanding. Conversely, tension and annoyance peaked significantly in the difficult condition. The co-occurrence of high Flow and high tension in the difficult condition suggests a dual state of engagement. In reflex-based games, tension does not necessarily break Flow; instead, it appears to be an inherent part of the high arousal demand. This implies that the optimal experience in rhythm games is not always a purely pleasant state, but rather one of intense, stressful engagement that players find rewarding despite the accompanying frustration.

5.5 Limitations and Future Directions

While this study provides insights into how reflex-based challenges affect Flow and game experience, several limitations should be acknowledged.

First, there are ongoing empirical concerns regarding the construct validity of the GEQ. Specifically, there is potential overlap between the Challenge, Tension/Annoyance, and Negative Affect constructs (Johnson et al., 2018). Furthermore, as argued by McDaniel et al. (2024), distinguishing between avatar skill and player skill is critical; since this study isolated player skill, reaction time becomes the primary variable. In this study, while these were treated as distinct subscales to address RQ3, the simultaneous peak of all three in the difficult condition suggests that players may struggle to distinguish between the high-intensity demand of the game and the resulting emotional frustration.

Second, although the within-subject design controlled for individual differences, the repetitive nature of *Geometry Dash* levels may have introduced practice effects or fatigue. Given the high-intensity attempts required to progress through the difficult condition, participants' reported Flow and negative affect might have been influenced by their increasing frustration with repetitive failure rather than the difficulty level alone.

Third, the sample predominantly included 26- to 33-year-olds with beginner to intermediate gaming abilities. This limits the generalisability of the findings to younger, older, or more expert populations. Expert-level gamers, for instance, may possess a higher enjoyability threshold, potentially experiencing lower levels of negative affect in difficult conditions compared to the participants in this study.

Fourth, Geometry Dash's single-input design supports the isolation of reflex responses but may not represent the more complex mechanics of multi-button combat or action-rhythm games.

Fifth, although the within-subject design was sufficient to identify main trends across difficulty conditions, the small sample may have limited the ability to detect subtle differences between the medium and difficult levels. The pilot study successfully informed refinement of the game levels before the main study, but a larger sample using the final level settings would likely provide greater statistical power for detecting subtle differences across closely aligned difficulty conditions.

Future research could address these constraints as follows:

Methodological Triangulation: Pairing the GEQ with physiological measures, such as EEG or heart-rate variability, could help resolve the overlap between subjective tension and objective challenge by providing neural correlates of the player's emotional state. Additionally, incorporating objective behavioural data, such as input precision and reaction time, can aid in upholding subjective Flow reports (Lora-Ariza et al., 2022).

Adaptive Designs: Future studies could employ Dynamic Difficulty Adjustment (DDA) (Lora-Ariza et al., 2022) to mitigate fatigue and practice effects, ensuring the challenge-skill balance remains

constant and reducing the risk of forced negative affect from repetitive failure.

Diverse Demographics and Power: Expanding recruitment to include professional or expert-level gamers would clarify whether the patterns of reflex challenge engagement and the resulting negative affect remain consistent across different skill tiers and age groups. Furthermore, expanding the total participant pool in future iterations would provide a larger participant sample necessary to firmly separate close performance variables at the upper limits of reflex-based challenge.

Mechanical and Genre Expansion: Future work should transition from minimalist single input mobile setups to more complex control architectures. Testing player Flow within multi-button arcade combat systems or fast-paced action rhythm-games would clarify if these reflex mechanics and emotional profiles remain consistent across multi-dimensional control schemes.

5.6 Implications and Future work

Acquiring a deeper understanding of how tactical challenge influences player engagement offers practical implications for game designers, particularly in crafting experiences that maintain Flow and engagement without narrative support. This is reinforced by Umar et al. (2022), who emphasise that identifying isolated immersive features is a robust method for modern game selection. Beyond entertainment, the cognitive engagement generated by reflex-based challenges has significant potential in health and education. As noted by Ottiger et al. (2021), Flow experiences are increasingly used in rehabilitation settings for patients with neurological diseases to encourage motor and cognitive recovery. Furthermore, in educational contexts, the link between action-oriented gameplay and enhanced visual processing, as demonstrated by Bhattachariyya et al. (2017), suggests that reflex-based mechanics can be a powerful tool for cognitive training in younger populations. By shifting the focus away from narrative or graphical immersion and toward the intensity of speed and time-precision, this research provides an initial framework for using mechanical demand as an important component of optimal experience in both recreational and therapeutic software.

6. CONCLUSION

This study examined how reflex-based challenges in *Geometry Dash* modulate Flow, sensory immersion, and affective states. The results indicate that Flow is significantly higher in medium-difficulty conditions, supporting Csikszentmihalyi's (1990) theory that a balance between demand and ability is necessary to enter the 'Flow channel'. Notably, the absence of a significant difference in Flow between the medium and difficult levels is consistent with a potential 'Flow plateau', although this interpretation is limited by the study's sample size and statistical power to detect subtle differences between these conditions. These findings suggest that the intense nature of reflex-based tactical challenge may help maintain engagement even at high levels of difficulty, provided the game properly increases difficulty to keep pace with player skill (McDaniel, Davis, and Weeks, 2024). This stability of engagement parallels the findings of Lora-Ariza et al. (2022), suggesting that once a specific threshold of behavioural control is achieved, the Flow state can be sustained even as mechanical intensity approaches the upper limits of player capability.

A key contribution of this work is the finding that Challenge was a consistent predictor of Flow across all conditions, whereas perceived Competence showed a more condition-dependent relationship with Flow. This suggests that at peak engagement, players may experience a loss of self-consciousness. In this state, reactive gameplay and the performative nature of movement (Blackman, 2024) override the conscious monitoring of one's own skill. Furthermore, the stability of sensory immersion across levels, aligned with findings by Jennett et al. (2008), indicates that in non-narrative environments, audiovisual absorption appears to operate relatively independently of difficulty-induced Flow. This supports the assertion that while the mastery of game challenges (Kuo et al., 2024) drives engrossment, certain immersive states can be decoupled from the challenge-skill balance in reflex-dependent environments. While limitations regarding sample homogeneity and potential practice effects remain, the study provides a foundation for future research. Utilising physiological measures and Dynamic Difficulty Adjustment (DDA) (Lora-Ariza et al., 2022) could further refine our understanding of these dynamics. Ultimately, this research emphasises that reflex-based engagement follows a unique pattern: one where Flow can coexist with tension, and where sensory immersion remains a constant backdrop to the fluctuating demands of tactical challenge. These insights offer practical guidance for developers seeking to balance frustration and enjoyment in high-speed, mechanically-driven games.

7. REFERENCES

Argasiński, J. K., & Węgrzyn, P. (2018, June 20). Affective Patterns in Serious Games. *Future Generation Computer Systems*, 526-538. Retrieved from https://doi.org/10.1016/j.future.2018.06.013

(2013). Retrieved from RobTop Games: https://www.robtopgames.com/

Adams, E. (2004, July 9). *Postmodernism and the Three Types of Immersion*. Retrieved from Gamasutra.com:

http://designersnotebook.com/Columns/063_Postmodernism/063_postmodernism.htm

Adams, E. (2014). *Fundamentals of Game Design* (3rd ed.).

Bhattachariyya, P., Das, S., & R., A. (2017). Exposure to Video Games Shortens Simple Visual Reaction Time: A Study in Indian School Children.

Blackman, T. (2024, December 13). Virtual reality and videogames: immersion, presence, and the performative spatiality of 'being there' in virtual worlds. *Social & Cultural Geography, 25*(3), 404-422. doi:10.1080/14649365.2022.2157041

Brown, E., & Cairns, P. (2004). A Grounded Investigation of Game Immersion. *ACM CHI* , (pp. 1297-1300). Vienna.

Cairns, P., Cox, A., & Nordin, A. I. (2014). Immersion in Digital Games: Review of Gaming Experiences. In M. C. Angelides, & H. Agius (Eds.), *Handbook of Digital Games* (pp. 339-361). John Wiley & Sons, Inc.

Canossa, A., Badler, J., Magy, S.-N., & Anderson, E. (2016, September). Eliciting Emotions in Design of Games – a Theory Driven Approach.

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental Methods: Between-Subject and Within-Subject Design. *Journal of Economic Behavior & Organization*, 1-8.

Chen, J. (2007, April). Flow in Games (and Everything Else). *COMMUNICATIONS OF THE ACM, 50*(4), 31-34.

Cox, A. L., Cairns, P., Shah, P., & Carroll, M. (2012). Not Doing But Thinking: The Role of Challenge in the Gaming Experience. *ACM CHI .* Texas.

Csikszentmihalyi, M. (1990). Flow: The Psychology of Optimal Experience.

Csikszentmihalyi, M., & Larson, R. (2014). The Experience Sampling Method. In M. Csikszentmihalyi, *Flow and the Foundations of Psychology.* Springer. doi:10.1007/978-94-017-9088-8_2

Denisova, A., Cairns, P., Guckelsberger, C., & Zendle, D. (2020). Measuring perceived challenge in digital games: Development & validation of the challenge originating from recent gameplay interaction scale (CORGIS). *International Journal of Human-Computer Studies*.

Ermi, L., & Mäyrä, F. (2005). Fundamental Components of the Gameplay Experience: Analysing Immersion. *Digital Games Research Association DiGRA.*

Gerling, K. M., Birk, M., Mandryk, R. L., & Doucette, A. (2013). The Effects of Graphical Fidelity on Player Experience. (pp. 229-236). Tampere: AcademicMindTrek.

Guillen, G., Jylhä, H., & Hassan, L. (2021). The Role Sound Plays in Games: A Thematic Literature Study on Immersion, Inclusivity and Accessibility in Game Sound Research.

Haehn, L., Böffel, C., & Schlittmeier, S. (2023). Exploring the Impact of Ambient and Character Sounds on Player Experience. doi:10.20944/preprints202312.0324.v1

IJsselsteijn, W. A., de Kort, Y., Poels, K., & Jurgelionis, A. (2007). Characterising and Measuring User Experiences in Digital Games. *ACE Conference'07 .* Salzburg, Austria: ACM.

IJsselstijn, W. A., de Kort, Y. A., & Poels, K. (2013). The Game Experience Questionnaire. FUGA The fun of gaming: Measuring the human experience of media enjoyment.

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and Defining the Experience of Immersion in Games. *International Journal of Human-Computer Studies, 66*, 641-661. doi:10.1016/j.ijhcs.2008.04.004

Johnson, D., Gardner, M. J., & Perry, R. (2018). Validation of two game experience scales: The Player Experience of Need Satisfaction (PENS) and Game Experience Questionnaire (GEQ). *International Journal of Human-Computer Studies*, 38-46. Retrieved from https://doi.org/10.1016/j.ijhcs.2018.05.003

Klasen, M., Weber, R., Kircher, T. T., Mathiak, K. A., & Mathiak, K. (2012). Neural Contributions to Flow Experience During Video Game Playing. *SCAN, 7*, 485-495. doi:10.1093/scan/nsr021

Kuo, C.-H., Chen, M.-J., Nabahan, R., & She, H.-C. (2024). Space Adventure Game-Based Learning: How Games and Scaffolds Affect Eigth Graders' Physics Learning and Game Immersion. *IEEE TRANSACTIONS ON LEARNING TECHNOLOGIES, 17*, 229-240. doi:10.1109/TLT.2023.3288879

Lakens, D. (2022). Sample Size Justification. *Collabra: Psychology, 8*(1). Retrieved from https://doi.org/10.1525/collabra.33267

Li, T., & Raferty, S. (n.d.). Playing Geometry Dash with Convolutional Neural Networks.

Lora-Ariza, D. S., González-Calero, P. A., Sánchez-Ruiz, A. A., & Camps-Ortueta, I. (2022, December). Measuring Control to Dynamically Induce Floow in Tetris. *IEEE Transactions on Games, 14*(4), 579-588.

Lu, A. S., & Moller, A. C. (2024). Elaborating the Role of Narrative and Self-Determination Theory in Video Games Research. *Interacting with Computers*, 1-15. Retrieved from https://doi.org/10.1093/iwc/iwae058

Maxwell, S. E., Delaney, H. D., & Kelley, K. (2017). *Designing Experiments and Analyzing Data: A Model Comparison Perspective, Third Edition* (3rd ed.). Routledge. Retrieved from https://doi.org/10.4324/9781315642956

McDaniel, E. L., Davis, J., & Weeks, M. (2024). Assessing the Flow State in a Classic Video

Game through Simulation. *SouthEastCon 2024* (pp. 1119-1124). IEEE. doi:10.1109/SOUTHEASTCON52093.2024.10500281

Mcmahan, A. (2003, January). Immersion, engagement, and presence: A method for analyzing 3-D video games.

Michailidis, L., Balaguer-Ballester, E., & He, X. (2018, September). Flow and Immersion in Video Games: The Aftermath of a Conceptual Challenge. *Frontiers in Psychology, 9*. doi:10.3389/fpsyg.2018.01682

Ottiger, B., Van Wegen, E., Keller, K., Nef, T., Nyffeler, T., Kwakkel, & Vanbellingen, T. (2021). Getting into a "Flow" state: a systematic review of flow experience in neurological diseases. *Journal of NeuroEngineering and Rehabilitation, 18*(65). Retrieved from https://doi.org/10.1186/s12984-021-00864-w

Procci, K., & Bowers, C. (2011). An Examination of Flow and Immersion in Games. *HUMAN FACTORS and ERGONOMICS SOCIETY 55th ANNUAL MEETING*, (pp. 2183-2187).

Rutrecht, H., Wittmann, M., Koshnoud, S., & Igarzábal, F. A. (2021, March 30). Time Speeds Up During Flow States: A Study in Virtual Reality. *Timing & Perception*, 353–376. doi:10.1163/22134468-bja10033

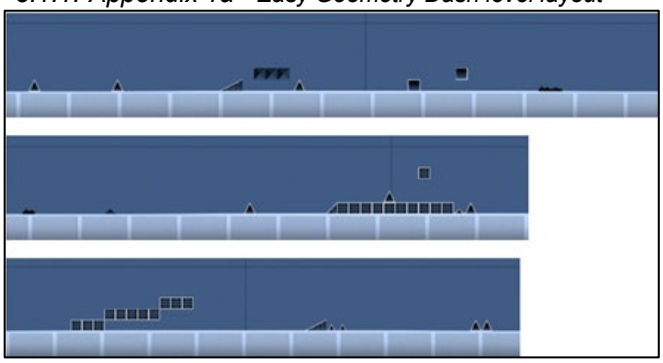
Sweetser, P., & Wyeth, P. (2005, July). GameFlow: A Model for Evaluating Player Enojoyment in Games. *ACM Computers in Entertainment, 3*(3).

Umar, N., Yuyun, & Gani, H. (2022, December 20). Generating game immersion features for immersive game selection. *ILKOM Jurnal Ilmiah, 14*(3), 264-274. Retrieved from http://dx.doi.org/10.33096/ilkom.v14i3.1224.264-274

8. APPENDICES

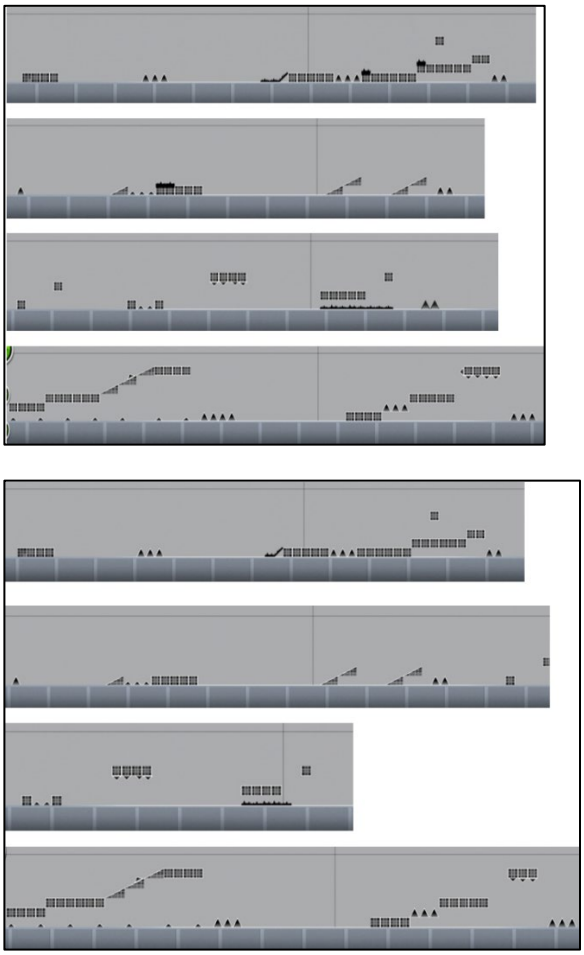
8.1 Appendix 1 – Game level designs: Easy, medium (before and after changes), and difficult (before and after changes)

8.1.1. Appendix 1a - Easy Geometry Dash level layout



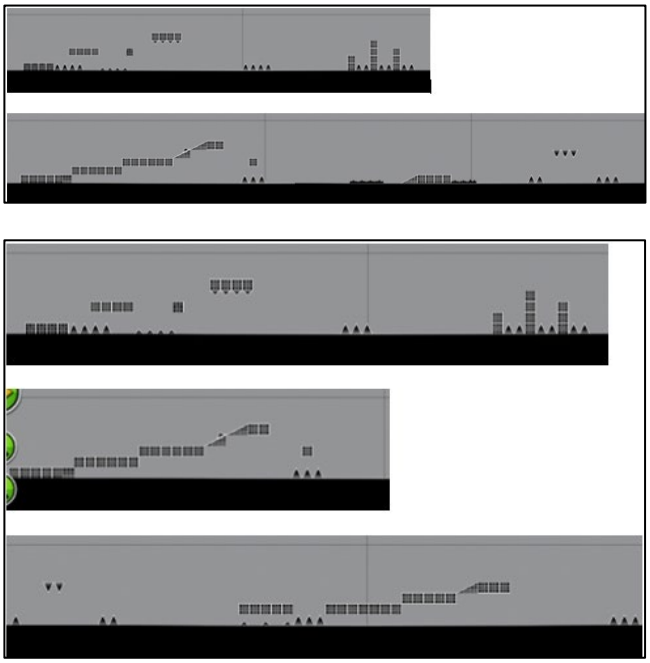
Note. Speed level 1×, 17 jumps

8.1.2. Appendix 1b - *Medium Geometry Dash level layout before and after adjustment*



Note. Image 1: Speed level 2×, 21 jumps; Image 2: Speed level 2×, 18 jumps

8.1.3. Appendix 1c - *Difficult Geometry Dash level layout before and after adjustment*



Note. Image 1: Speed level 3×, 17 jumps; Image 2: Speed level 3×, 22 jumps

8.2 Appendix 2 – Imaginary Immersion, Positive and Negative Affect, and Combined Immersion Repeated-Measures ANOVA Results

8.2.1. Appendix 2a - Imaginary immersion

Imaginary immersion scores showed no significant effect of difficulty level, $F(2,34) = 1.76$, $p = .187$, $\eta^2 = .027$, with sphericity assumed ($W = 0.901$, $p = .435$). All pairwise comparisons were non-significant (all ps > .05), indicating stable imaginary immersion across Easy ($M_{Easy} = 1.2$, $SD = 1.1$), Medium ($M_{Medium} = 1.3$, $SD = 1.0$), and Difficult ($M_{Difficult} = 0.9$, $SD = 0.9$) conditions. However, all conditions showed significant non-normality (all ps < .05), and model residuals significantly deviated from normality ($W = 0.922$, $p = .002$), strongly suggesting the need for non-parametric verification.

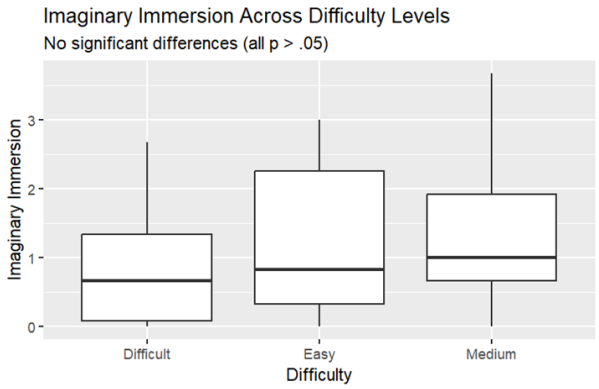


Figure 11: *Imaginary immersion by difficulty level*

8.2.2. *Appendix 2b - Negative affect*

Negative affect showed no significant variation across difficulty levels, $F(2,34) = 0.43$, $p = .654$, $\eta^2 = .006$, with sphericity assumed ($W = 0.844$, $p = .257$). All pairwise comparisons were non-significant (all ps = 1.000), indicating stable negative affect in Easy ($M_{Easy} = 0.8$, $SD = 0.8$), Medium ($M_{Medium} = 0.7$, $SD = 0.7$), and Difficult ($M_{Difficult} = 0.9$, $SD = 0.8$) conditions. However, all conditions showed severe non-normality (all ps ≤ .010), and model residuals significantly deviated from normality ($W = 0.866$, $p < .001$), requiring non-parametric verification.

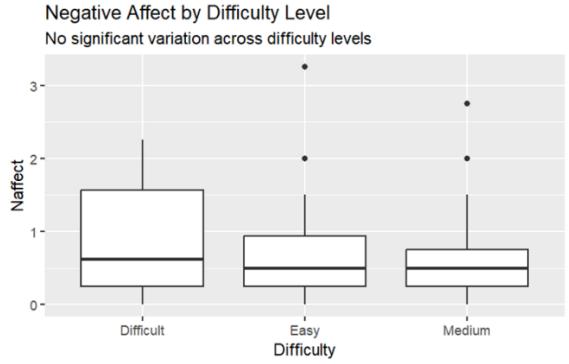


Figure 13: *Negative affect remained consistent across all difficulty levels*

8.2.3. Appendix 2d - Combine immersion (sensory and imaginary)

The combined sensory-imaginary immersion analysis showed no significant effect of difficulty level, $F(2,34) = 1.40$, $p = .260$, $\eta^2 = .016$, with sphericity assumed ($W = 0.785$, $p = .145$). All pairwise comparisons were non-significant (all ps > .05), indicating stable immersion across Easy ($M_{Easy} = 1.6$, $SD = 1.0$), Medium ($M_{Medium} = 1.7$, $SD = 1.0$), and Difficult ($M_{Difficult} = 1.4$, $SD = 0.9$) conditions. However, the Medium condition showed marginal non-normality ($p = .037$), and model residuals significantly deviated from normality ($W = 0.927$, $p = .003$), suggesting caution in interpreting these parametric results.

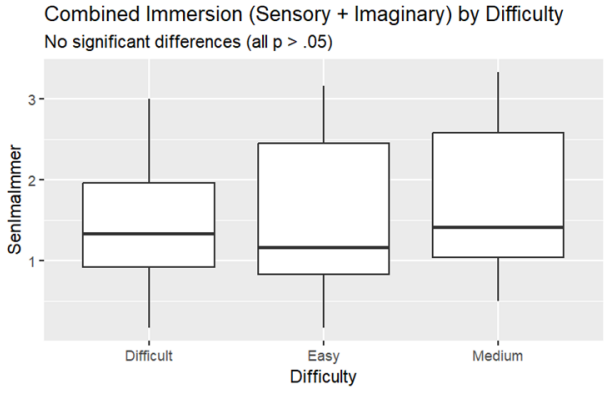


Figure 14: *Immersion levels remained stable across all task difficulties*

8.3 Appendix 3 – Assumption and Multicollinearity Tests for All Levels

Assumption tests for all difficulty levels (Easy, Medium, Difficult) confirmed regression models were robust (see Figures 1 – 9). For clarification, all correlations used here are based on Pearson's r. $p < .05$, and the analysis used the means of Challenge, Competence, and Flow

constructs. Residuals showed mild non-normality (left-skew in Medium/Difficult) but no heteroscedasticity was found. Predictor independence was upheld, with Challenge-Competence correlations ranging from r = -0.446 (Easy) to r = -0.353 (Difficult), all below multicollinearity thresholds (VIFs < 1.5). Challenge consistently correlated strongly with Flow (Easy: r = 0.740; Medium: r = 0.548; Difficult: r = 0.532), while Competence-Flow relationships varied from negligible (Medium: r = 0.010) to moderate (Difficult: r = 0.410).

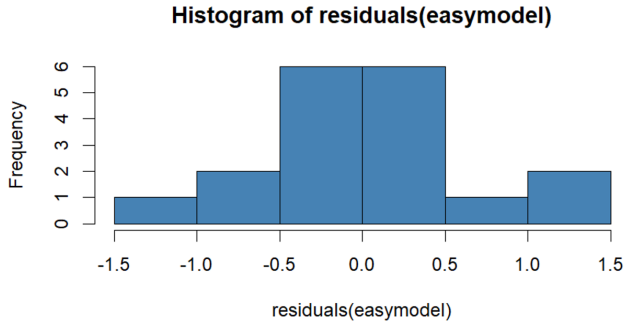


Figure 15: *Histogram suggests approximate normality, slightly right-skewed but not severe deviations. residuals vs. fitted plot shows homoscedasticity (no funnel shape)*

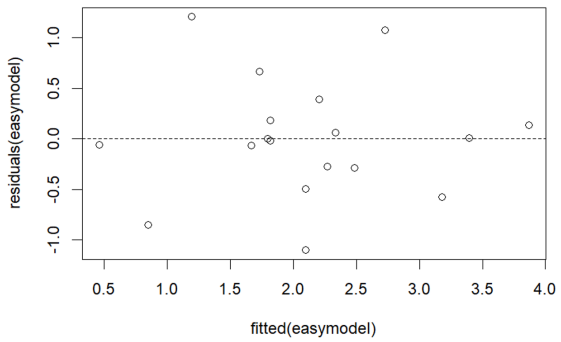


Figure 16: *Residuals vs. fitted plot for the easy difficulty shows homoscedasticity (no funnel shape)*

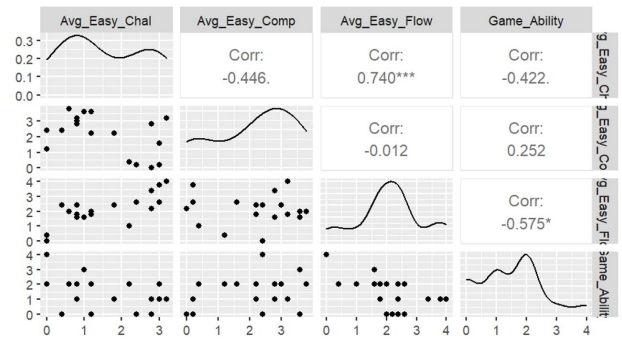


Figure 17: *Multicollinearity diagnostics revealed optimal conditions for regression analysis at Easy level. The moderate negative correlation between Challenge and Competence (r = -0.446) ensures predictor independence, while Challenge's strong correlation with Flow (r = 0.740) confirms its primary predictive role. The negligible Competence-Flow correlation (r = -0.012) suggests its effect emerges only when considered alongside Challenge. Additionally, a moderate negative relationship shows up between Game Ability and Flow (r = -0.575)*

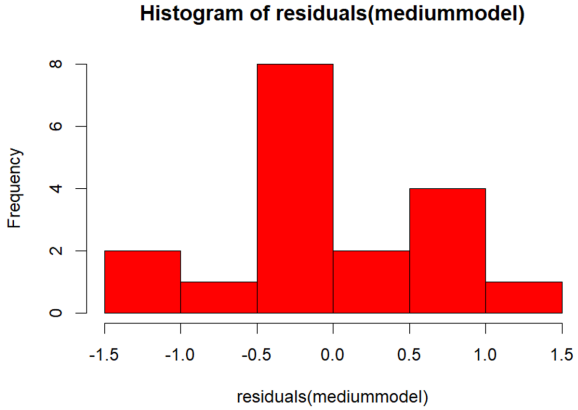


Figure 18: *Slight left-skewness (cluster near zero) suggests mild non-normality, but the model remains robust given regression's tolerance to moderate deviations*

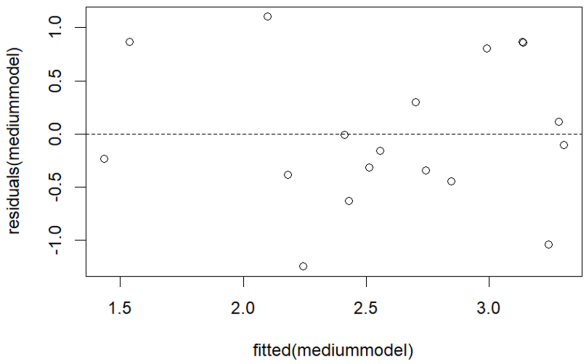


Figure 19: *Residuals vs. fitted plot for the medium level shows random scatter around zero with no funnel pattern confirms homoscedasticity. Minor outliers (e.g., residuals ±1.5) do not violate regression assumptions*

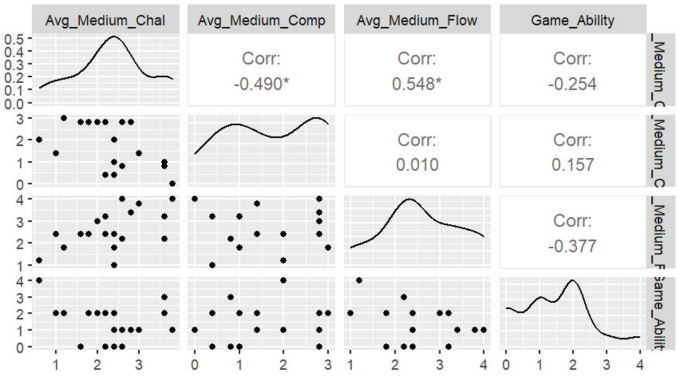


Figure 20: *Challenge-Flow shows moderate correlation (r = 0.548) at Medium level, while Competence-Flow is negligible (r = 0.010). Predictors, Challenge-Competence, exhibit acceptable independence (r = -0.490 VIFs < 1.5), ruling out multicollinearity. Additionally, there was no significant correlation between Game Ability and the other*

constructs

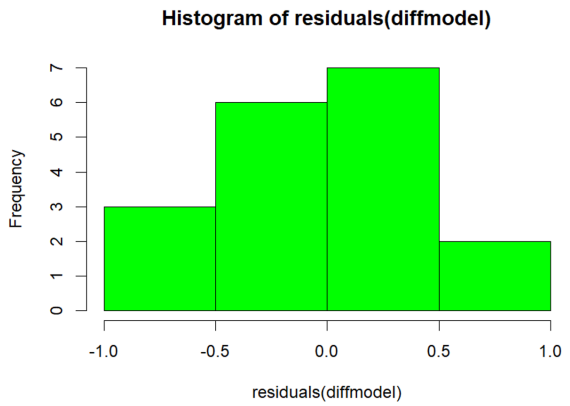


Figure 21: *Residuals cluster near zero with moderate left skew (-1.0 to -0.5). While non-ideal, the regression remains valid given its robustness to mild non-normality*

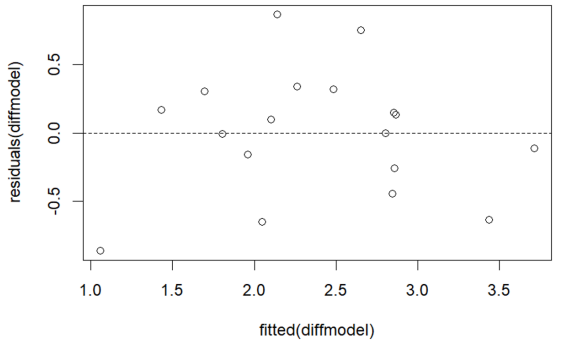


Figure 22: *Residuals vs. fitted plot for the difficult level shows random dispersion around zero confirms homoscedasticity. No systematic patterns suggest the model accounts for variance appropriately despite one outlier (~-1.5)*

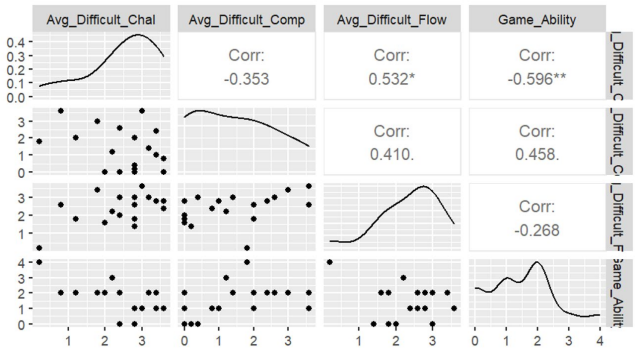


Figure 23: *Challenge-Flow correlation is significant (r = 0.532), while Competence-Flow is moderate but non-significant (r = 0.410). Predictors show acceptable independence (r = -0.353). Additionally, Game Ability and Challenge exhibit a strong negative correlation (r = -0.596**)*