HOW ADAPTIVE NON-PLAYER CHARACTERS (NPCS) SHAPE PLAYER CHALLENGE AND IMMERSION

Group 1H – Yuxu Ge, Anand Mohan, Luke Jackson, Nithish Sagar Dharmappa, Rohan Hanumanthappa Channagouder

University of York

WHY DOES THIS TOPIC MATTER?

CHALLENGE PERCEPTION FROM PLAYERS IN DIGITAL GAMES

Players expect challenge to feel fair, reactive, and believable

Unpredictable or illogical NPCs break immersion

NPC behaviour shapes how challenge is perceived

RESEARCH QUESTION:

How do players experience NPCs that adapt to their play style and skill level, and which AI architectures or behaviour models allow for this adaptation without reducing immersion or narrative coherence?

Examines how NPC adaptation influences players' perception of challenge

Clarifies the key concepts involved: adaptation, immersion, and narrative logic

Sets up the need to analyse existing work across adaptive AI, PCG, and challenge-measurement research

PERSONA-DRIVEN PROCEDURAL CONTENT GENERATION

Fernandes et al. (2021) [4]: Procedural Content Generation(PCG) – The study used the four personas and three gameplay scenarios to generate the personalized levels

Example from the study:
In the game "Grave Rave", for the PCG levels and enemy behavior the level generator was dependent on the input

Four Personas(Rule based agents):

- R01(The Shooter)
- R02(The Defender)
- R03(The Hunter Grave Robber)
- R04(The Mad Man)

Gameplay scenarios:

- Hardcore
- Close call
- Easy

Therefore: It has strong internal validity as it follows rule-based actions but limited external and ecological validity as agents are not real players, and tested in a simple game environment, the link to the perceived challenge of players is limited.

BELIEVABILITY & IMMERSION

- *Butt et al. (2024) [5]:* ANFIS ('Neuro-Fuzzy' systems) can be used to train NPCs to imitate human players (with an average accuracy of 95%)
- Improves trust, believability and immersion
- Low ecological validity due to simple testing environments

ADAPTIVE NPC BEHAVIOUR

Overview

Research in this cluster focuses on enabling NPCs to adjust their behaviour dynamically in response to player actions, aiming to improve engagement and maintain challenge.

Representative Work

Supli & Xu (2025) [1]: Deep RL NPCs → 73% player preference

Shows improved engagement through learned adaptive behaviour.

Liu et al. (2025) [2]: Hybrid RL + Behaviour Trees

Combines adaptability with predictable, designer-controlled logic.

Common Limitation

These approaches often emphasise behavioural control but pay limited attention to how such adaptation affects immersion or narrative coherence.

MEASURING THE CHALLENGE

Overview

This cluster focuses on how players perceived challenge is measured, aiming to capture the psychological components that shape their gameplay experience.

Representative Work

Denisova et al. (2020) [3]: CORGIS scale

Identifies four dimensions of challenge: performative, cognitive, emotional, and decision-related.

Demonstrates strong construct validity for measuring players' challenge experience.

Common Limitation

Challenge-measurement tools operate independently from adaptive AI or PCG research, meaning subjective challenge is rarely linked to NPC behaviour or dynamic content systems.

SYNTHESIS

Adaptive NPC Behaviour

- Enables responsive and dynamic behaviour
- Focuses on control, decision-making, and challenge adjustment
- *Rarely considers immersion or narrative coherence*

Difficulty & Content Adaptation

- Adjusts challenge through level or content changes
- Uses player models/personas for tuning
- Rarely evaluates players' subjective challenge perception

Challenge Measurement

- Provides validated tools to assess how players experience challenge
- Captures cognitive, emotional, and performative aspects
- *Not linked to AI or content adaptation systems*

Overall Synthesis

Together, these clusters describe how challenge is created (NPC / content) and how challenge is experienced (measurement) - but they operate in isolation, without informing each other.

EVIDENCE GAP

1. Lack of integration across the three domains
Existing research treats NPC adaptation, content adaptation, and challenge-measurement as separate problems, without examining how they interact.

2. No model connecting adaptive behaviour with perceived challenge
We know how to adapt NPCs and generate adaptive content, but we do not know how these adaptations affect players' subjective experience of challenge or immersion.

3. Immersion and narrative logic are rarely considered
Many adaptive systems optimise difficulty, but few consider whether adaptation feels believable or breaks narrative coherence.

Therefore:
There is no unified approach that explains how NPCs can adapt to players while preserving immersion and narrative logic, directly motivating our RQ.

Deep Dive

Denisova, Cairns, Guckelsberger & Zendle (2020)

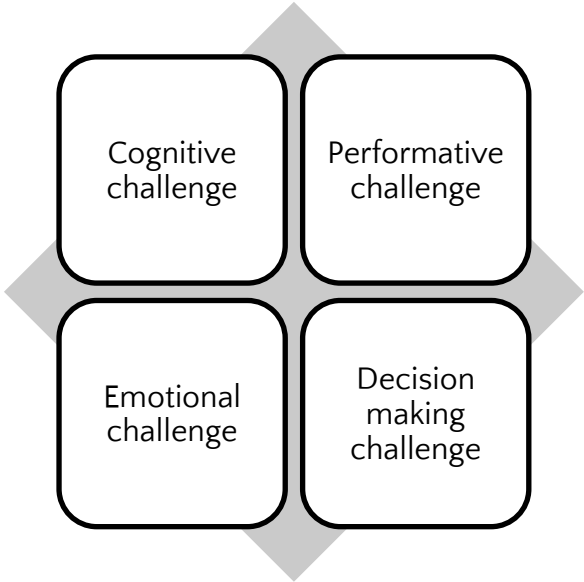
Measuring perceived challenge in digital games: Development & validation of the CORGIS scale

Study Overview

This study developed and validated *CORGIS*, a psychometric scale for measuring perceived challenge in digital games.

Researchers surveyed literature, conducted interviews with players, created an item pool, and validated the final scale with over 1400 players across two studies.

The scale identifies four challenge dimensions:



Deep Dive

Denisova, Cairns, Guckelsberger & Zendle (2020)

Measuring perceived challenge in digital games: Development & validation of the CORGIS scale

Key Validity Issues

Construct Validity

- Challenge was conceptualised via literature + interviews
- Four factors emerged with clear theoretical meaning
- But "Decision-making challenge" overlapped with emotional challenge to limited discriminant validity

Internal Validity

- Two-stage survey design (EFA to CFA) ensures internal reliability
- Large diverse samples (N:400 + N:1000)
- Item selection based on IRT, improving precision
- But self-report immediately after gameplay → possible recall bias

External Validity

- Validated on three games: *Life is Strange*, *XCOM*, *Monster Hunter World*
- Challenge patterns aligned with genre expectations
- But mostly Western players recruited via Reddit/Twitter to sample bias

Ecological Validity

- Measured right after actual gameplay
- Captures multi-dimensional challenge experiences
- But no real-time in-game behavioural data to relies purely on subjective responses

Deep Dive

Denisova, Cairns, Guckelsberger & Zendle (2020)

Measuring perceived challenge in digital games: Development & validation of the CORGIS scale

Why This Matters

- Provides a standardised, validated instrument to measure challenge across games
- Enables researchers to separate different challenge types instead of treating “difficulty” as a single variable
- Supports better experimental control (e.g., isolating cognitive vs performative challenge)
- Useful for game designers to evaluate whether their intended challenge structure is perceived as expected

Deep Dive

Denisova, Cairns, Guckelsberger & Zendle (2020)

Measuring perceived challenge in digital games: Development & validation of the CORGIS scale

Proposed Improvements

1. **Add gameplay telemetry –** Combine subjective CORGIS scores with in-game logs (failures, reaction time, decision latency) to improve ecological validity.

2. **Improve discriminant validity –** Decision-making and emotional challenge strongly correlate; clarifying item wording or separating moral-dilemma constructs could help.

3. **Cross-cultural validation –** Validate CORGIS with players from non-Western regions to ensure generalisability.

4. **Test sensitivity across difficulty modes –** Study whether CORGIS scores change reliably under Easy / Normal / Hard settings.

5. **Short-form scale –** A 10-item short version could reduce participant fatigue in industry applications.

Conclusion

- Adaptive NPC behaviour and content adjustment influence how players *experience* challenge, but only when these adaptations remain consistent with the game's narrative logic and do not appear artificial.
- Immersion is not only about difficulty balance: players remain immersed when adaptive behaviour feels fair, predictable within the story world, and aligned with how human opponents would act.
- Tools like CORGIS provide the missing subjective layer, allowing researchers to evaluate challenge perception rather than relying solely on performance metrics or system-level adjustments.

Overall Insight

Across the literature, challenge is created by NPC behaviour and game structure, but *experienced* through cognitive, emotional, and performative responses – yet these layers are rarely studied together.

Evidence Gap Restated

No existing study integrates adaptive AI, adaptive content, and validated measures of challenge perception. Future work must bridge these layers to understand whether adaptation genuinely enhances challenge and immersion.

References

[1] A. A. Supli and X. Siqi, “Leveraging deep reinforcement learning for dynamic NPC behavior and enhanced player experience in unity3d,” Entertainment Computing, vol. 55, p. 101007, Jul. 2025, doi: https://doi.org/10.1016/j.entcom.2025.101007.

[2] T. Liu, A. Cann, I. Colbert, and M. Saeedi, “Combining Reinforcement Learning and Behavior Trees for NPCs in Video Games with AMD Schola,” arXiv.org, 2025. https://arxiv.org/abs/2510.14154.

[3] A. Denisova, P. Cairns, C. Guckelsberger, and D. Zendle, “Measuring perceived challenge in digital games: Development & validation of the challenge originating from recent gameplay interaction scale (CORGIS),” International Journal of Human-Computer Studies, vol. 137, p. 102383, May 2020, doi: https://doi.org/10.1016/j.ijhcs.2019.102383.

[4] P. M. Fernandes, J. Jorgensen, and Niels N. T. G. Poldervaart, “Adapting Procedural Content Generation to Player Personas Through Evolution,” 2021 IEEE Symposium Series on Computational Intelligence (SSCI), pp. 01–09, Dec. 2021, doi: https://doi.org/10.1109/ssci50451.2021.9659880.

[5] N. A. Butt, Mian Muhammad Awais, Samra Shahzadi, T. Kim, and I. Ashraf, “Towards the development of believable agents: Adopting neural architectures and adaptive neuro-fuzzy inference system via playback of human traces,” Journal of King Saud University - Computer and Information Sciences, vol. 36, no. 8, pp. 102182–102182, Sep. 2024, doi: https://doi.org/10.1016/j.jksuci.2024.102182.

Questions?