Acta Ludologica | 2022, Vol. 5, No. 1

Performance-Based Dynamic Difficulty Adjustment and Player Experience in a 2D Digital Game: A Controlled Experiment

Nigel Robb, Bo Zhang

Nigel Robb, BA MA PhD. MSc.
Hokkaido University
Research Faculty of Media and Communication
Kita 17, Nishi 8, Kita-ku
060-0817 Sapporo, Hokkaido
JAPAN
nigelrobb@imc.hokudai.ac.jp

Nigel Robb is a Specially Appointed Associate Professor in the Research Faculty of Media and Communication at Hokkaido University in Sapporo, Japan. He holds a PhD in philosophy and an MSc in software engineering, both from Queen's University Belfast in Northern Ireland. At Hokkaido University, he teaches practical English language skills to undergraduate and graduate students, as well as courses on video games and other digital media.

Bo Zhang, BSc. MSc. PhD.
East China Normal University
Faculty of Education
North Zhongshan Road 3663, Putuo District
2000 63 Shanghai
CHINA
bzhang@ed.ecnu.edu.cn

Bo Zhang is an Assistant Researcher in the Department of Education Information Technology, Faculty of Education at East China Normal University. She received her PhD in Inclusive Design and Creative Technology Innovation from University College Dublin in Ireland. Her research interests are using virtual worlds, virtual reality, augmented reality and ICT in Education, CSCL, and Inclusive Design. Her current study is to explore the application of markerless AR cognitive training to help children with autism to improve their cognitive abilities in China.

ABSTRACT:
Dynamic difficulty adjustment (DDA) in digital games involves altering the difficulty of a game based on real-time feedback from the player. Some approaches to DDA use measurements of player performance, such as success rate or score. Such performance-based DDA systems aim to provide a bespoke level of challenge to each player, so that the game is neither too hard nor too easy. Previous research on performance-based DDA shows that it is linked to better player performance but finds mixed results in terms of player experience (e.g., enjoyment). Also, while the concept of flow is regarded as an important aspect of digital game experience, little research has considered the effects of performance-based DDA on flow. We conducted an experiment on the effects of performance-based DDA on player performance, enjoyment, and experience of flow in a digital game. 221 participants played either the DDA version of the game, a control version (difficulty remained constant), or an incremental version (difficulty increased regardless of performance). Results show that the DDA group performed significantly better. However, there were no significant differences in terms of enjoyment or experience of flow.

KEY WORDS:
adaptive software, digital games, dynamic difficulty adjustment, flow, game balancing, performance.

Introduction

Most digital games involve some challenge for the player. Some games are generally easy, some are generally hard, and almost all games feature some change in the difficulty level over time; typically, games get harder the further the player progresses. In this paper, we refer to this traditional approach as 'incremental difficulty adjustment'. Furthermore, a diverse range of people (e.g., in terms of age, gender, motivations, and preferences) play games.[1] Taken together, these points begin to illustrate the complex challenges involved for the game designer when determining the difficulty of a game, to enhance the experience for a range of players. If the game is too easy, more skilled players may be bored; but if it is too hard, less skilled players may be frustrated.[2] It is likely that this applies regardless of the genre or intended purpose of the game. Whether it is a fast-paced action game intended to entertain, a puzzle game for mathematics education, a cognitive training game for children with cognitive impairments, or even a language-learning application with game-like features, it is obviously essential that players engage optimally with the software to ensure the desired outcome. As such, the level of challenge provided by a game is an important consideration.

Within this broad issue, one potential solution to some of these challenges lies in dynamic difficulty adjustment (DDA). DDA in digital games refers to any technique in which the difficulty of the game is altered during the game (or perhaps between games) in

1 PIERRE-LOUIS, S.: *2021 Essential Facts About the Video Game Industry*. 2021. [online]. [2022-05-13]. Available at: <https://www.theesa.com/wp-content/uploads/2021/08/2021-Essential-Facts-About-the-Video-Game-Industry-1.pdf>.

2 LEIKER, A. M. et al.: The Effects of Autonomous Difficulty Selection on Engagement, Motivation, and Learning in a Motion-Controlled Video Game Task. In *Human Movement Science*, 2016, Vol. 49, No. 5, p. 327.

response to some feedback about the player's experience.[3] The aim is to continuously tailor the difficulty of the game to each individual player. DDA has been featured in digital games since at least 1981.[4] The promise of DDA lies in the fact that the game designer does not need to pre-determine one specific difficulty curve (or even a range of pre-determined curves, as in games which let the player select, e.g., Easy, Medium, or Hard mode). Instead, the designer can effectively provide a range of possible difficulties which are dynamically selected for each individual player based on their experience of the game. Essentially, this provides a bespoke difficulty curve for each player.

Approaches to Dynamic Difficulty Adjustment

One important distinction when using DDA is based on the metric used to provide feedback for game adaptation. In this paper, we have so far (intentionally) characterised this feedback very broadly, as 'player experience'. In practice, player experience can be determined by various means. We can broadly categorise approaches to DDA in two ways, depending on whether they use players' in-game performance to provide the feedback (performance-based DDA), or use some information about the player's affective state (affective DDA). A combination of these approaches would of course also be possible.

Performance-based DDA involves measuring the player's performance in the game and adjusting the level of challenge provided accordingly. Previous research on performance-based approaches to DDA demonstrates the wide range of choice available in the design of such systems, both in terms of the indicator of player skill (i.e., the feedback), and the game features that are subsequently adjusted. Regarding the measurement of player skill, previous approaches have used, for example, the time taken to complete a task,[5] players' scores,[6] or, in more complex systems, multiple measurements may be combined and evaluated to determine the current state of the player.[7] Regarding the game features that are adjusted, these range from simple adjustments such as changing the speed and layout of the game[8] or changing the number and kind of objects a player must interact with,[9] to more complex systems which alter the behaviour or characteristics of

3 ZOHAIB, M.: Dynamic Difficulty Adjustment (DDA) in Computer Games: A Review. In *Advances in Human-Computer Interaction*, 2018, Vol. 11, No. 1, p. 2. [online]. [2022-05-13]. Available at: <https://www.hindawi.com/journals/ahci/2018/5681652/>.

4 ADAMS, E.: *The Designer's Notebook: Difficulty Modes and Dynamic Difficulty Adjustment*. Released on 14th May 2008. [online]. [2022-05-13]. Available at: <https://www.gamasutra.com/view/feature/132061/the_designers_notebook_.php>.

5 SHAREK, D., WIEBE, E.: Investigating Real-time Predictors of Engagement: Implications for Adaptive Videogames and Online Training. In *International Journal of Gaming and Computer-Mediated Simulations*, 2015, Vol. 7, No. 1, p. 26.

6 BATEMAN, S. et al.: Target Assistance for Subtly Balancing Competitive Play. In TAN, D. (ed.): *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York : ACM, 2011, p. 2358.

7 HUNICKE, R.: The Case for Dynamic Difficulty Adjustment in Games. In LEE, N. (ed.): *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology*. New York : ACM, 2005, p. 430-431.

8 ALEXANDER, J. T., SEAR, J., OIKONOMOU, A.: An Investigation of the Effects of Game Difficulty on Player Enjoyment. In *Entertainment Computing*, 2013, Vol. 4, No. 1, p. 55.

9 ROBB, N., WALLER, A., WOODCOCK, K. A.: Developing a Task Switching Training Game for Children with a Rare Genetic Syndrome Linked to Intellectual Disability. In *Simulation & Gaming*, 2019, Vol. 50, No. 2, p. 171.

computer-controlled enemy characters[10] or dynamically generate the layout of the game environment.[11]

Affective DDA refers to any approach which aims to use some feedback about the player's emotional state as the basis for the adaptivity. A growing body of research has investigated the feasibility of using psychophysiological measurement to provide an index of players' emotions during gameplay and adapt the game experience based on these measurements. Measures used include cardiovascular data (e.g., heart rate), electro-dermal activity (i.e., galvanic skin response, which is directly dependent on sweat-gland activity), electromyography (which measures muscle movements), and neuroimaging techniques. The latter category includes techniques such as electroencephalography and functional near-infrared spectroscopy; both of which use sensors attached to the head to provide real-time measurements of brain activity with a relatively high temporal resolution.[12] These techniques have been used to adapt game difficulty, for example, by increasing the speed of the game or decreasing the size of targets. In addition, some studies have used affective feedback to alter other features of a game not related to difficulty, such as lighting and audio effects. For a comprehensive review of research in this area and references for all examples discussed in this paragraph, see B. Bontchev.[13] One obvious disadvantage of affective-based DDA is the requirement for additional equipment (which is often large and expensive) to obtain the psychophysiological feedback. Therefore, affective DDA is most likely not yet suitable for widespread use in digital games. However, technological advances will undoubtedly address this issue. For example, preliminary work shows the potential of using machine learning techniques and a standard video camera for remote, non-contact detection of player emotions.[14] However, due to this current limitation of affective DDA, we will focus on performance-based DDA in the remainder of this paper.

Effects of Performance-Based Dynamic Difficulty Adjustment

Previous research has investigated the effects of using performance-based DDA on various aspects of the game playing experience. Several studies in this area have considered the relationship between DDA and player enjoyment. J. T. Alexander et al. used an experimental game to investigate how DDA compared with incremental difficulty adjustment.[15]

10 HUNICKE, R.: The Case for Dynamic Difficulty Adjustment in Games. In LEE, N. (ed.): *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology*. New York : ACM, 2005, p. 431.

11 SHAKER, N., YANNAKAKIS, G., TOGELIUS, J.: Towards Automatic Personalized Content Generation for Platform Games. In YOUNGBLOOD, G. M., BULITKO, V. (eds.): *Sixth Artificial Intelligence and Interactive Digital Entertainment Conference*. Burnaby : PKP Publishing Services, 2010, p. 65. [online]. [2022-05-13]. Available at: <https://ojs.aaai.org/index.php/AIIDE/article/view/12399>.

12 THIBAULT, R. T., LIFSHITZ, M., RAZ, A.: The Self-Regulating Brain and Neurofeedback: Experimental Science and Clinical Promise. In *Cortex*, 2016, Vol. 74, No. 1, p. 249.

13 For more information, see: BONTCHEV, B.: Adaptation in Affective Video Games: A Literature Review. In *Cybernetics and Information Technologies*, 2016, Vol. 16, No. 3, p. 3-34.

14 BEVILACQUA, F., ENGSTRÖM, H., BACKLUND, P.: Game-Calibrated and User-Tailored Remote Detection of Stress and Boredom in Games. In Sensors, 2019, Vol. 19, No. 13, p. 3. [online]. [2022-05-13]. Available at: <https://www.mdpi.com/1424-8220/19/13/2877/htm>.

15 ALEXANDER, J. T., SEAR, J., OIKONOMOU, A.: An Investigation of the Effects of Game Difficulty on Player Enjoyment. In *Entertainment Computing*, 2013, Vol. 4, No. 1, p. 53.

They found that, while casual gamers reported enjoying a simple 2D game more when the difficulty was dynamically adjusted according to their performance, experienced gamers (who made up most of the sample) enjoyed the game more when the difficulty was adjusted incrementally. This study also showed that players enjoyed the game more when the difficulty was tailored to their gaming experience (casual v experienced) rather than their performance.[16] However, in a sample of 90 players, only 19 were categorised as casual players. Furthermore, this classification was determined by the players' response to the question "are you a casual or experienced gamer?".[17] It is therefore difficult to determine how to understand the distinction between casual and experienced gamers as the classification criteria are not explicit.

A. Nagle et al. also found that performance-based DDA led to lower player enjoyment than an alternative system in which players could control the level of difficulty throughout the game themselves. They created a game in which players had to memorize a list of objects and locations, then find the objects and place them in the correct location. The difficulty of the game was determined by the number of objects (more is harder) and the number of times they could view the list of objects and location numbers (fewer is harder). However, while player enjoyment was lower with DDA, DDA was associated with better player performance. That is, when they allowed players to control the number of objects and number of times the list could be consulted, performance was significantly lower than when these values were automatically adjusted based on player performance.[18]

D. Sharek and E. Wiebe also showed that DDA was associated with better player performance.[19] Using an isometric puzzle game, they created over 100 different levels which were tested and ordered by difficulty. They had three difficulty conditions: DDA, in which more difficult levels were provided to players based primarily on measures of performance; incremental; and a choice condition, in which players were given the option to select a harder or easier level after each completed level. Players in the DDA condition showed significantly higher performance (indexed as reaching more difficult levels more quickly) than players in the other two conditions.[20] However, in a study by K. A. Orvis et al., no significant differences in performance or motivation were found between 4 groups,[21] playing versions of a game with either no difficulty adjustment, incremental difficulty adjustment, or adaptive adjustment (two versions, distinguished in terms of the starting difficulty, which was either easy or hard).[22] Other research on motivation finds similar negative results, with DDA not associated with significantly different levels of player motivation than incremental difficulty adjustment in a Spanish language education game.[23] However, in line with previous results showing increased performance, the authors showed that DDA led to significantly higher learning outcomes, which they attribute to a scaffolding effect wherein the reductions in difficulty (the "scaffold") are provided when students need support, then removed when students were ready to progress.

16 ALEXANDER, J. T., SEAR, J., OIKONOMOU, A.: An Investigation of the Effects of Game Difficulty on Player Enjoyment. In *Entertainment Computing*, 2013, Vol. 4, No. 1, p. 59-60.
17 Ibidem, p. 56.
18 NAGLE, A. et al.: The Effect of Different Difficulty Adaptation Strategies on Enjoyment and Performance in a Serious Game for Memory Training. In CLUA, E., VILAÇA, J. (eds.): *IEEE 3rd International Conference on Serious Games and Applications for Health*. Rio de Janeiro : IEEE, 2014, p. 127-132.
19 SHAREK, D., WIEBE, E.: Investigating Real-time Predictors of Engagement: Implications for Adaptive Videogames and Online Training. In *International Journal of Gaming and Computer-Mediated Simulations*, 2015, Vol. 7, No. 1, p. 20.
20 Ibidem, p. 25-28.
21 ORVIS, K. A., HORN, D. B., BELANICH, J.: The Roles of Task Difficulty and Prior Videogame Experience on Performance and Motivation in Instructional Videogames. In *Computers in Human Behavior*, 2008, Vol. 24, No. 5, p. 2424.
22 Ibidem, p. 2420.
23 SAMPAYO-VARGAS, S. et al.: The Effectiveness of Adaptive Difficulty Adjustments on Students' Motivation and Learning in an Educational Computer Game. In *Computers & Education*, 2013, Vol. 69, No. 10, p. 452.

D. Altimira et al. found that DDA increased player engagement in a digitally augmented game of table tennis. By projecting images onto a surface, they could increase or decrease the difficulty of the game (e.g., by altering the size of the virtual table projected onto the surface). They found that adjusting the difficulty in response to the score differentials between the two players (e.g., by making one player's half of the table smaller, thus making the game harder for the opposing player), was associated with significantly higher player engagement (self-report questionnaire) than no adjustment.[24] Preliminary results from by research by S. Xue et al. using DDA in a mobile game distributed via the Google Play Store and Apple App Store show that DDA increases player engagement over a longer period (4 months). This study is notable as it measures player engagement objectively, in terms of total time spent playing the game. The authors also note that using DDA had no effect on the amount of revenue generated from in-game transactions.[25]

D. Altimira et al. also highlight another potential application of DDA technology, in that they showed that using DDA significantly reduced the score differences between players, thus allowing less skilled players to be more competitive against more skilled players.[26] Other studies have successfully used DDA to reduce skill differentials between players of different abilities.[27] K. M. Gerling et al. created a rhythm game (i.e., in which players must perform steps in time with music) which could either be controlled by a dance mat (i.e., the player inputs the rhythm with their feet), buttons on a standard game controller, or a via a wheelchair input (wheelchair movements were captured by a motion sensor camera). Using various techniques, they produced adaptive versions of the game which allowed less-skilled able-bodied players to compete with more-skilled able-bodied players, and players with and without mobility disabilities to play the game together.[28] S. Bateman et al. also decreased performance differentials between players by providing adaptive targeting assistance in a simple shooting game.[29] DDA may therefore be important for enabling people with disabilities to play multiplayer games with those without disabilities.[30] DDA may also be one factor which can increase the effectiveness of rehabilitation games for people with disabilities, both in terms of making such games accessible and in terms of adapting the difficulty of the games to suit players of a wide range of abilities and provide an optimum level of challenge, which is shown to increase the effectiveness of such games.[31]

To summarise, previous research generally supports the idea that performance-based DDA can increase player performance, and that it can be used to reduce performance differentials between players of different abilities. Proposed applications of this

---

24 ALTIMIRA, D. et al.: Enhancing Player Engagement Through Game Balancing in Digitally Augmented Physical Games. In *International Journal of Human-Computer Studies*, 2017, Vol. 103, No. 7, p. 35-42.

25 XUE, S. et al.: Dynamic Difficulty Adjustment for Maximized Engagement in Digital Games. In BARRET, R., CUMMINGS, R. (eds.): *WWW '17 Companion: Proceedings of the 26th International Conference on World Wide Web Companion*. Geneva : International World Wide Web Conferences Steering Committee, 2017, p. 470.

26 ALTIMIRA, D. et al.: Enhancing Player Engagement Through Game Balancing in Digitally Augmented Physical Games. In *International Journal of Human-Computer Studies*, 2017, Vol. 103, No. 7, p. 43.

27 HWANG, S. et al.: How Game Balancing Affects Play: Player Adaptation in an Exergame for Children with Cerebral Palsy. In MIVAL, O. (ed.): *Proceedings of the 2017 Conference on Designing Interactive Systems*. New York : ACM, 2017, p. 704-705.

28 GERLING, K. M. et al.: Effects of Balancing for Physical Abilities on Player Performance, Experience and Self-Esteem in Exergames. In JONES, M., PALANQUE, P. (eds.): *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*. New York : ACM, 2014, p. 2203-2207.

29 BATEMAN, S. et al.: Target Assistance for Subtly Balancing Competitive Play. In TAN, D. (ed.): *Proceedings of the SIGCHI Conference on Human Factors in Computing System*. New York : ACM, 2011, p. 2363.

30 HERNANDEZ, H. A. et al.: Designing Action-Based Exergames for Children with Cerebral Palsy. In MACKAY, W. (ed.): *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York : ACM, 2013, p. 1269.

31 BARRETT, N. et al.: The Use and Effect of Video Game Design Theory in the Creation of Game-Based Systems for Upper Limb Stroke Rehabilitation. In *Journal of Rehabilitation and Assistive Technologies Engineering*, 2016, Vol. 3, No. 1, p. 7. [online]. [2022-05-13]. Available at: <https://journals.sagepub.com/doi/full/10.1177/2055668316643644>.

include making games more accessible or inclusive for people with disabilities, and, related to this, increasing the effectiveness of games for rehabilitation. However, in terms of player experience – i.e., engagement, enjoyment, and motivation – previous research provides mixed results on the effects of DDA.

Flow

Several issues relevant to DDA are captured in the concept of flow, which was first introduced by M. Csikszentmihalyi in the 1970s. During a flow state, an individual is completely focused on an activity; it is an enjoyable and fulfilling experience (an 'optimal' experience), often described colloquially as being in "the Zone".[32] Csikszentmihalyi identifies several characteristics of the flow state, including having clear goals, concentrating on the task at hand, feeling in control, and being engaged in a challenging activity requiring skill. Flow was originally modelled by M. Csikszentmihalyi as an optimal balance between anxiety and boredom; the zone in which one is challenged enough to not be bored, but skilled enough to not be anxious about one's performance.[33] Since M. Csikszentmihalyi's original work, a large body of research has further investigated and characterised flow, and some of this work has explicitly focused on digital games, which have been claimed to "possess ideal characteristics to create and maintain flow experiences in that the flow experience of video games is brought on when the skills of the player match the difficulty of the game".[34] The importance of the flow experience is shown by empirical research suggesting that flow is a source of digital games' appeal to players in the long-term, and the flow experience predicts players' intention to play games.[35] In educational games, flow has been used as a measure of game quality, and a small amount of research suggests that the experience of flow may be associated with the effectiveness of game-based learning.[36]

The notion of a balance between player skill and game difficulty shows the direct relevance of flow to performance-based DDA. If the aim of these approaches to DDA is to match the difficulty of a game to each unique player's skill level, then it seems likely that DDA could be used to achieve a balance between these two factors, and therefore encourage flow experiences. However, although theoretical discussions of flow feature in much research on DDA, few studies have investigated the relationship between difficulty adjustment and flow empirically.

In one such study, D. Ang and Mitchell showed that playing with incremental difficulty adjustment and with a version of DDA in which the player could control the difficulty were both associated with significantly different scores[37] (compared to no DDA) on several constructs of the Flow State Scale.[38] This included challenge-skill balance, which was greater when participants played a game with DDA. A second study by the same authors investigated how player experience (including experience of flow) was affected by

32 CHEN, J.: Flow in Games (And Everything Else). In *Communications of the ACM*, 2007, Vol. 50, No. 4, p. 31.
33 CSIKSZENTMIHALYI, M.: The Flow Experience and Its Significance for Human Psychology. In CSIKSZENTMIHALYI, M., CSIKSZENTMIHALYI, I. S. (eds.): *Optimal Experience: Psychological Studies of Flow in Consciousness*. Cambridge : Cambridge University Press, 1988, p. 30.
34 SHERRY, J. L.: Flow and Media Enjoyment. In *Communication Theory*, 2004, Vol. 14, No. 4, p. 340.
35 PERTTULA, A. et al.: Flow Experience in Game Based Learning – A Systematic Literature Review. In *International Journal of Serious Games*, 2017, Vol. 4, No. 1, p. 58.
36 Ibidem, p. 67-68.
37 ANG, D., MITCHELL, A.: Comparing Effects of Dynamic Difficulty Adjustment Systems on Video Game Experience. In SCHOUTEN, B., MARKOPOULOS, P., TOUPS, Z. (eds.): *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. New York : ACM, 2017, p. 321.
38 See also: JACKSON, S. A., MARSH, H.: Development and Validation of a Scale to Measure Optimal Experience: The Flow State Scale. In *Journal of Sport & Exercise Psychology*, 1996, Vol. 18, No. 1, p. 17-35.

changes in how and when players were able to adjust the difficulty of a game, although this study did not use a no-DDA control condition.[39] Most importantly, the DDA used in both these studies is not performance-based as discussed in this paper, in that players manually (and voluntarily) increased or decreased the difficulty during the game themselves, rather than have the difficulty automatically adjusted based on a measurement of player performance.

The Present Study

Our aim in this study was to investigate how a performance-based approach to DDA affects player performance, enjoyment, and experience of flow, using an experimental game created for this study. We conducted a controlled experiment with 3 groups, with each group playing a different version of the game. The independent variable was the way in which the difficulty of the game was adjusted, with 3 levels: (1) DDA, (2) incremental difficulty adjustment (in which the game gets progressively harder irrespective of player experience), and (3) no difficulty adjustment (control group). In the DDA version of the game, we used a simple algorithm to adjust the difficulty, and two measures of player performance provided the feedback upon which the adjustments were based. Our hypotheses were:

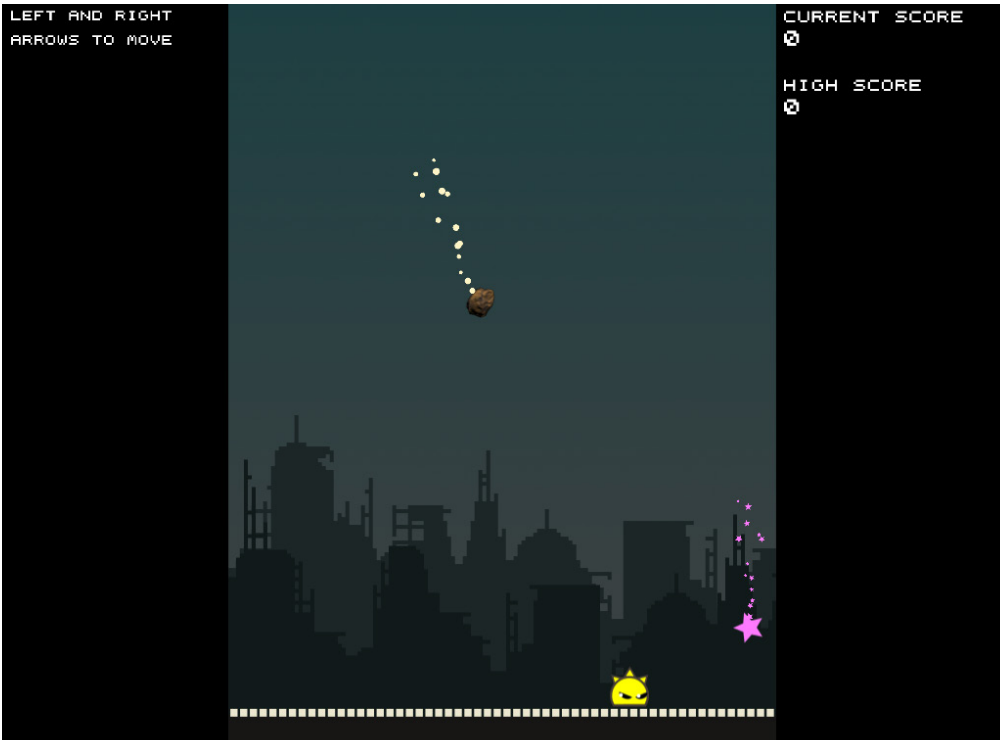
- H1: DDA will produce greater player performance than either incremental or no difficulty adjustment.
- H2: DDA will lead to greater experience of flow than either incremental or no difficulty adjustment.
- H3: DDA will lead to greater enjoyment than either incremental or no difficulty adjustment.

Methodology

To test our hypotheses, we designed and implemented a simple digital game called *Meteor Shower* (Picture 1). The object of the game is to avoid the meteors which continually fall from the top of the screen while catching the pink falling stars. The player controls the yellow character by moving left or right along the bottom of the screen (using the left and right directional arrows on the keyboard). The velocity of the meteors determines the difficulty of the game, with higher velocities making the game more difficult (i.e., the meteors are harder to avoid). The game consists of 20 levels, each lasting 45 seconds, with a short pause between each level. Players are awarded one point for each star they catch and lose a point each time a meteor hits the character. The score is reset to 0 at the end of each level. There are 8 falling stars to catch in each level (each falling 5 seconds apart), and so the maximum score available on any level is 8 (i.e., the player catches all stars and avoids all meteors). While the meteors and stars appear to the player to originate from random locations, the game in fact uses a seeded random number generator to ensure that the pattern of locations at which stars and meteors appear is the same for each player. When generated, stars fall in a straight line. Meteors move in a straight line

39 ANG, D., MITCHELL, A.: Representation and Frequency of Player Choice in Player-Oriented Dynamic Difficulty Adjustment Systems. In ARNEDO, J., NACKE, L. E. (eds.): *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. New York : ACM, 2019, p. 589.

from their point of origin to the location of the player-controlled character at the time the meteor was generated. This ensures that players always must move the character to avoid every meteor.



Picture 1: The experimental game, Meteor Shower
Source: own processing

Three versions of the game were created. The sole difference between each version is the way in which the difficulty (i.e., the velocity of the meteors) was adapted during gameplay. In the control version of the game, the velocity of the meteors remained constant at 800 throughout. In the incremental difficulty adjustment version of the game, the velocity of the meteors increased by 50 at the start of each new level. The velocity on level 1 was 200; the velocity on level 20 was 1150. Finally, the DDA version of the game used a simple algorithm to adapt the velocity of the meteors in response to measurements of player performance. The starting velocity on level 1 was 900. The possible velocity settings ranged from 200 to 1700, increasing in increments of 50. Due to an error in the programming, value 1200 was not used.

After every 5 seconds of gameplay, two separate measures of the player's current success rate were calculated. The first of these measured the player's success at avoiding meteors over the previous 5 seconds of play (success1 = (nM - nH)/nM, where nM is the number of meteors in the previous 5 seconds and nH is the number of times the player was hit by a meteor in the previous 5 seconds). The second measurement was the player's overall success rate based on their score and the potential maximum score possible at that point in the level (success2 = score/nS, where nS is the number of stars that have fallen so far in the level). If either of these success rates was less than the target success rate (1; see following paragraph), the current setting of the game was judged to be too hard for the player and the velocity was therefore decremented (e.g., if the velocity was 350 it was reduced to 300), otherwise the current setting was judged to be too easy and the velocity was therefore incremented (e.g., if the velocity was 1400 it was increased to 1450. If the velocity was already at the lowest (200) or highest (1700) setting, no change was made.

Determining which value to use for the target success rate (i.e., the value at which the game was judged to be too easy), proved to be one of the most challenging design decisions in the development of the game, and we were unable to find previous research to guide this decision. Therefore, during the development process, we played versions of the game using success rates ranging from 0.75 (i.e., 75%) to 1 (i.e., 100%). We determined that the most satisfying experience was provided when we used a target success rate of 1 for both measures. Flow was measured using the Flow Short Scale which was first published in German[40] and later, in an English translation.[41] We used the online version of the scale[42] which has 14 items. Items 1-10 measure flow, items 11-13 measure anxiety, and item 14 measures perceived skill demands (challenge). In addition, demographic data (age, gender, frequency of digital game play) were collected.

During gameplay, each participant's score was recorded for levels 1-19. Due to a bug in the software (which we identified after running the experiment), the score for level 20 was not recorded. For the DDA group only, we also recorded the velocity of the meteors at 5 second intervals (i.e., each time the velocity was updated). This provided a measure of the difficulty of the game (higher velocity is more difficult), and a measure of the player's skill level (better players will reach higher velocities). Scores were recorded for all participants in the DDA group for 855 seconds of gameplay (i.e., not the full 15 minutes, due to a technical problem or bug, currently unidentified). We did not record velocity for the control group, as this remained constant throughout the game (800) or for the incremental group, as this increased on a predetermined scale with each level.

We conducted the experiment online using Amazon Mechanical Turk (MTurk). MTurk has been described as a "marketplace for work that requires human intelligence".[43] Users of the site are classified as "requesters" (who post tasks to the site) and "workers" (who complete these tasks in return for payment). Example tasks include completing surveys to provide feedback about a website, classifying images based on their content, or providing translations of short pieces of text. Typically, MTurk is appropriate for tasks which can be completed quickly, and for which many instances of the task must be completed. MTurk is now frequently used to conduct research in psychology[44] and it has been used in at least one previous study on the effects of DDA in digital games.[45]

However, several issues have been identified which may threaten the validity of data obtained from MTurk.[46] Some of these issues are not unique to MTurk. For example, the issue of selection bias, in the sense that MTurk workers choose the tasks they wish to complete, applies in any research wherein participants voluntarily opt to take part after

40 RHEINBERG, F., VOLLMEYER, R., ENGESER, S.: Die Erfassung des Flow-Erlebens. In STIENSMEIER-PELSTER, J., RHEINBERG, F. (eds.): *Diagnostik von Motivation und Selbstkonzept: Tests und Trends der pädagogisch-psychologischen Diagnostik – Band 2.* Göttingen : Hogrefe, 2003, p. 267-270.

41 ENGESER, S., RHEINBERG, F.: Flow, Performance and Moderators of Challenge-Skill Balance. In *Motivation and Emotion*, 2008, Vol. 32, No. 3, p. 170.

42 RHEINBERG, F.: *Flow Short Scale*. [online]. [2022-05-13]. Available at: <http://www.psych.uni-potsdam.de/people/rheinberg/messverfahren/fks1-e.html>.

43 ROUSE, S. V.: A Reliability Analysis of Mechanical Turk Data. In *Computers in Human Behavior*, 2015, Vol. 43, No. 2, p. 304.

44 CRUMP, M. J., McDONNELL, J. V., GURECKIS, T. M.: Evaluating Amazon's Mechanical Turk as a Tool for Experimental Behavioral Research. In *PLOS One*, 2013, Vol. 8, No. 3, p. 1. [online]. [2022-05-13]. Available at: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0057410>.

45 SHAREK, D., WIEBE, E.: Investigating Real-time Predictors of Engagement: Implications for Adaptive Videogames and Online Training. In *International Journal of Gaming and Computer-Mediated Simulations*, 2015, Vol. 7, No. 1, p. 27-28.

46 CHEUNG, J. H. et al.: Amazon Mechanical Turk in Organizational Psychology: An Evaluation and Practical Recommendations. In *Journal of Business and Psychology*, 2017, Vol. 32, No. 4, p. 349.

viewing, e.g., a poster or advertisement on social media. However, all participants in research conducted via MTurk will (obviously) have self-selected to become MTurk workers. Related to this, Cheung et al. note that samples obtained from MTurk may not be representative of the population of interest (e.g., all MTurk participants are internet users, which may not be appropriate for some studies).[47] On this issue, we make two observations. Firstly, we note that, in some respects, samples obtained from MTurk may be more diverse than samples obtained by traditional means. K. Casler et al. point out that most participants in psychological research are American college students; they showed that a sample of MTurk workers was significantly, desirably more diverse in terms of ethnicity, economic status, and age, than a sample of undergraduate students.[48] Secondly, we point out that the nature of our study – in which participants are required to play an online game remotely – dictates that participants would be required to have internet access and be reasonably computer literate whether recruited through MTurk or not.

Perhaps the most important concern with MTurk data is the possibility of participants answering questions without paying attention to the content (either fully, or at all, i.e., selecting random answers). However, steps can be taken to mitigate this risk.[49] Firstly, MTurk incorporates a rating system for workers, so that requesters can specify that only workers of a suitable quality can access their tasks. We will discuss this further in the Results section, where we describe how we used this system to specify that only workers of a certain quality could access our experiment. Secondly, items can be included in questionnaires to check for attentiveness (e.g., a multiple-choice item that states which option the respondent should select). Finally, it may also be possible to detect inattentive responses by analysing data gathered, although this would presumably depend on the nature of the data. The screening process we used to identify inattentive participants in the present study is described in the Results section.

The task was made available using MTurk's Survey Link template. This provides a link to an external website, where participants complete a task (typically a questionnaire) and receive a completion code. They then enter the completion code in MTurk to receive credit for completing the task. The default configuration of the Survey Link template allows each worker to only complete the task once. In our case, the survey link took participants to a site hosting the game. Which version of the game was loaded was determined randomly by the software. Participants pressed a button to start the game when they were ready. After 15 minutes of play, the game automatically ended, and participants were presented with a link to the webpage containing the questionnaire. When they submitted the questionnaire with all questions completed, the data were stored on a server, and a unique completion code was generated on the server and returned to participants. The completion code, a record of which version of the game they had played, and performance data automatically recorded during gameplay, were also stored on the server. Participants then entered the completion code in MTurk and were paid $0.99. All participants were paid, whether their data were included in the analysis or not.

Initially, we ran a pilot with 10 participants. By considering participants' performance data, it appeared that some participants did not actually play the game. It would be possible to merely let the game run for 15 minutes, then select random answers to the questions. To address this, we included two attention check items in the questionnaire. These

---

47 CHEUNG, J. H. et al.: Amazon Mechanical Turk in Organizational Psychology: An Evaluation and Practical Recommendations. In *Journal of Business and Psychology*, 2017, Vol. 32, No. 4, p. 349.

48 CASLER, K., BICKEL, L., HACKETT, E.: Separate but Equal? A Comparison of Participants and Data Gathered via Amazon's MTurk, Social Media, and Face-to-Face Behavioral Testing. In *Computers in Human Behavior*, 2013, Vol, 29, No. 6, p. 2158.

49 FLEISCHER, A., MEAD, A. D., HUANG, J.: Inattentive Responding in MTurk and Other Online Samples. In *Industrial and Organizational Psychology*, 2015, Vol. 8, No. 2, p. 199-200.

items stated that the participant should select option 7 ("very much") in the Likert-style scale. We also screened the data collected during the full experiment and removed data which appeared to show no engagement with the game (see Results). In addition, we decided to use MTurk's qualifications feature to ensure that the task was only available to workers who (1) had completed over 10000 tasks on MTurk; and (2) had an approval rate of 97% or higher.

Data were collected from 300 adults via MTurk; each randomly assigned to play either the control, DDA, or incremental versions of the game. Entries in which answers to either of the attention check questions were incorrect were removed. We also removed entries in which a score of zero was recorded for every level, as this suggested that the participants did not actually play the game, but merely let it run for the required time. Finally, we considered the velocity data recorded for participants in the DDA group and removed any entries in which the pattern of velocities across the levels suggested that participants had not actually played the game (i.e., when the velocity quickly decreased to 200 and did not rise above 250 for the remainder of the game). This left 221 participants (93 female) whose data were retained for analysis. Ages ranged from 20 years to 65 years, with a mean age of 36.51 years (std. deviation 9.73). There were 82 participants in the control group, 68 in the DDA group, and 71 in the incremental group.

Results

The 14-item online version of the Flow Short Scale was shown to have acceptable reliability (Cronbach's a = .834). Chi-square tests of homogeneity showed that the three groups did not significantly differ in terms of gender ($\chi2(3) = 1.4$, $p = .497$), number of days per week spent playing digital games ($\chi2(3) = 23.348$, $p = .055$), or age ($\chi2(3) = 69.525$, $p = .902$).

To analyse group differences in terms of overall mean score (i.e., participants' mean score over 19 levels of play), we ran a Kruskal-Wallis H test. The distributions of overall mean score were not similar for all groups. The DDA group had a smaller range (2.89) and smaller interquartile range (.83) than both the control group (range = 7.32, interquartile range = 3.01) and incremental group (range = 7.74, interquartile range = 2.89) (see Graph 1). The median values increased from the incremental group (3.74) to the control group (4.61) to the DDA group (5.05). These differences in median values were statistically significantly different between the groups, $\chi2(3) = 16.148$, $p < .0005$. Pairwise comparisons were then performed using Dunn's procedure with a Bonferroni correction for multiple comparisons. This analysis revealed statistically significant differences (adjusted p-values presented) in median values between the DDA group (mean rank = 135.31) and control group (mean rank = 106.91) ($p = .02$), and between the DDA group and incremental group (mean rank = 92.44) ($p < .0005$) groups, but not between the control group and the incremental group ($p = .488$).

We analysed the differences in scores between the three groups further by conducting one-way Welch ANOVAs on mean score per level (i.e., mean score for each group for each of 19 levels of play). A small number of outliers in the DDA group and Incremental group were not removed. Levene's test for equality of variances showed that the assumption of homogeneity of variances was violated in levels 4-19 (p-values ranging from .012 to <.0005). Scores were significantly different between the groups during levels 1, 3, 5, and levels 9-19. Games-Howell post hoc analyses revealed that the DDA group had higher mean scores than the control group in levels 3-19; these differences were significant in levels 3, 5, and 9-19 (p-values ranging from .019 to <.0005). The DDA group also had significantly higher mean scores than the incremental group in levels 10-19 (p-values ranging from .013 to <.0005). The control group had significantly higher mean scores than the incremental group in levels 14-19 (p-values ranging from .014 to <.0005). These results are shown in Charts 2, 3 and 4.

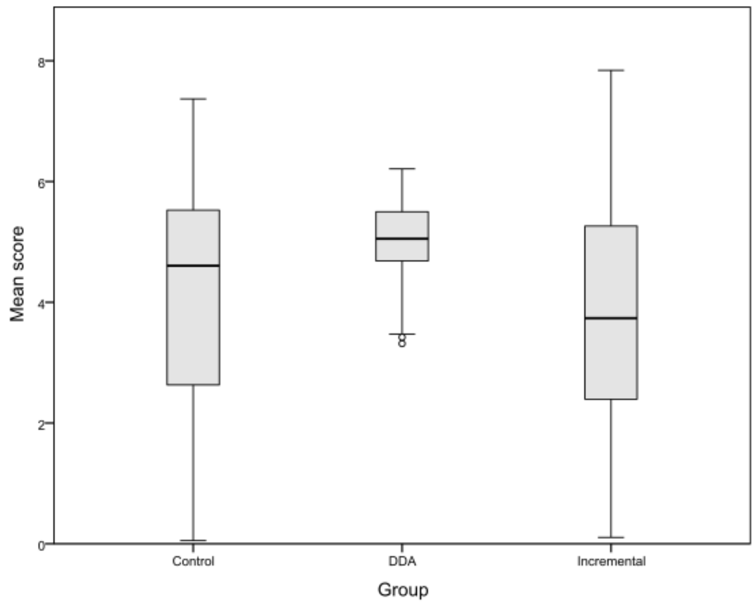


Chart 1: Boxplots of overall score for each of three groups – Control, Dynamic Difficulty Adjustment (DDA), and Incremental
Source: own processing

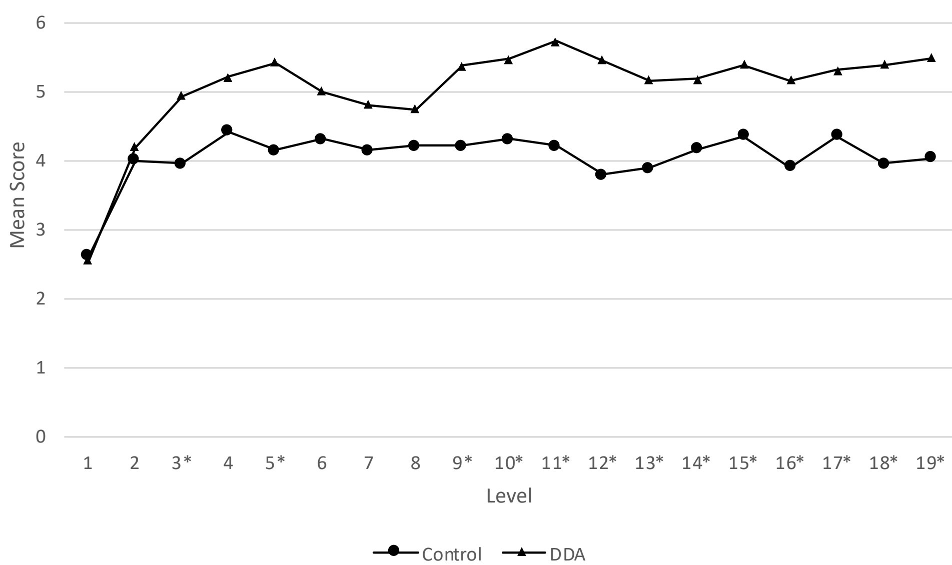


*Chart 2: A comparison of mean score per level for the Control and Dynamic Difficulty Adjustment (DDA) groups (significant differences are marked with a *)*
Source: own processing

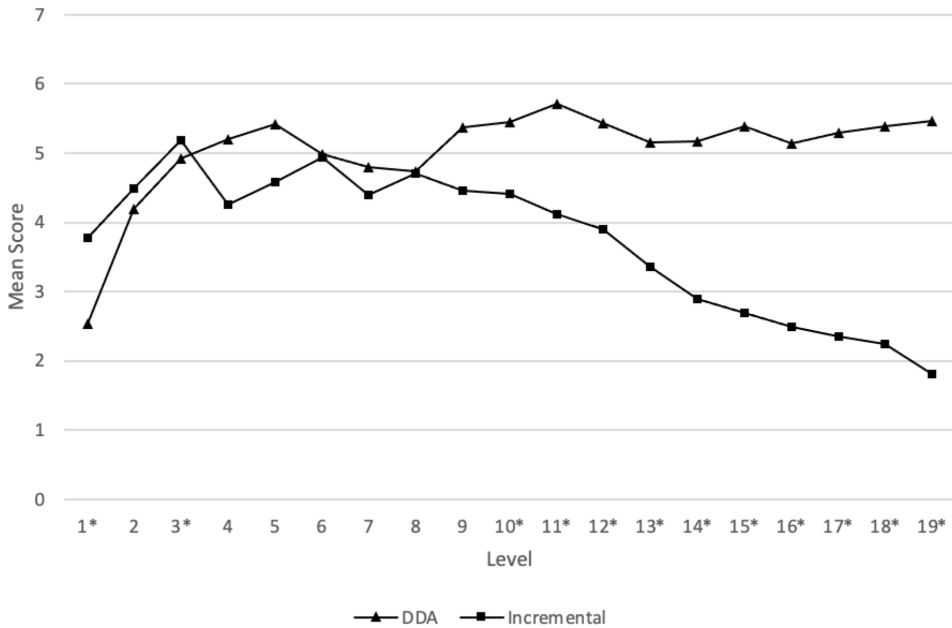


*Chart 3: A comparison of mean score per level for the Dynamic Difficulty Adjustment (DDA) and Incremental groups (significant differences are marked with a *)*
Source: own processing

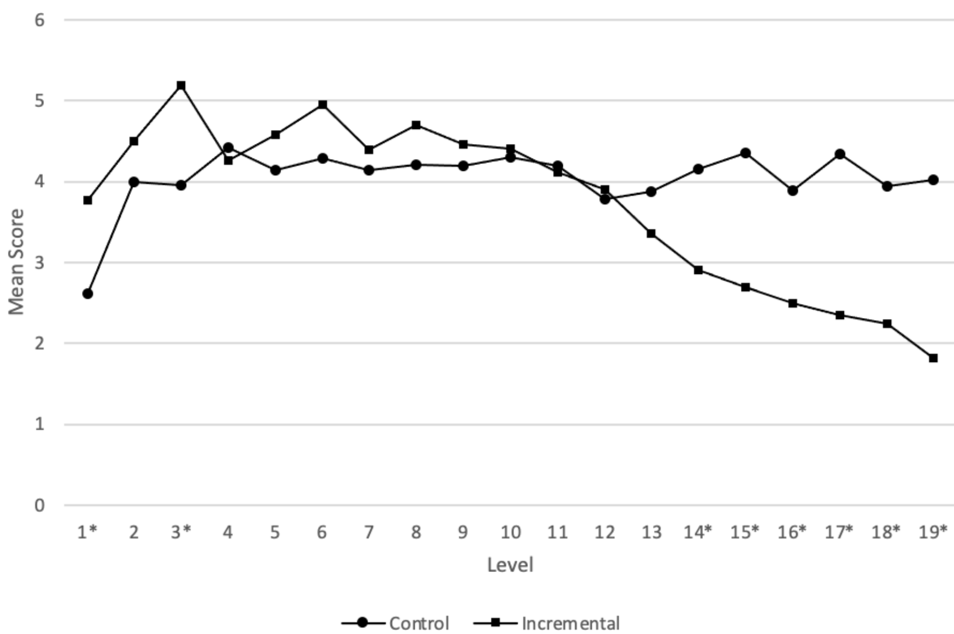


*Chart 4: A comparison of mean score per level for the Control and Incremental groups (significant differences are marked with a *)*
Source: own processing

For the DDA group, we analysed the velocity values, which indicates the difficulty of the game (higher velocity is more difficult). First, we considered the mean velocity across participants at each measurement point (5 seconds between each measurement, 171 measurements considered). Chart 4 shows how velocity ranged across participants over time. We considered the relationship between each participant's (DDA group only) overall mean velocity over 855 seconds of gameplay, and each participant's overall mean score across 19 levels, using Pearson's correlation test. The data were linear, overall mean score was normally distributed ($p > .05$), while overall mean velocity was not normally distributed ($p < .0005$). There was a strong positive correlation between overall mean score and overall velocity, $r(68) = .554$, $p < .0005$, with overall mean velocity explaining 31% of the variation in overall mean score. It means, as a higher velocity makes the game more difficult, and higher velocities are only achieved by players who perform better, velocity here can be used as an index of player performance.

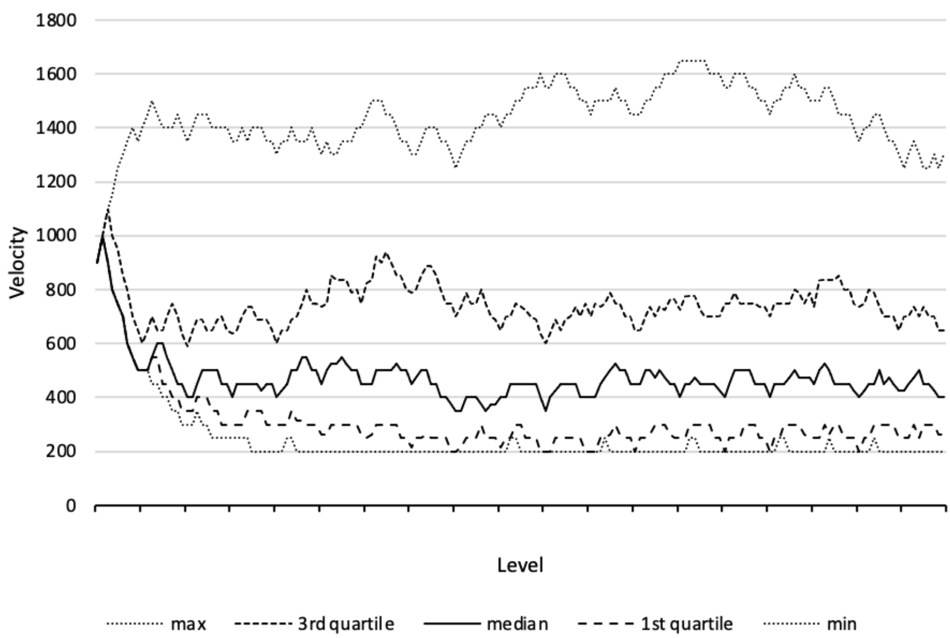


Chart 5: Range of meteor velocity per level for the Dynamic Difficulty Adjustment (DDA) group over 19 levels
Source: own processing

We also ran Kruskal-Wallis H tests to analyse differences between the groups in terms of the 3 factors of the Flow Short Scale, and the single-item for Enjoyment. In the case of Flow, the distributions were not similar for all groups (assessed by visual inspection of a box plot). The median values increased from the incremental group (4.9) to the DDA group (4.95), to the control group (5.3). These differences were not statistically significantly different between the groups, $\chi 2(3) = 4.087$, $p = .130$. In the case of Anxiety, the distributions were similar for all groups. The median values increased from the incremental group (3.67) to the control group (4.0) and DDA group (4.0). These differences were not statistically significantly different, $\chi 2(3) = 3.336$, $p = .189$. In the case of Challenge, the distributions were not similar for all groups and the median values were the same for all three groups (4.0). In the case of Enjoyment, the distributions were not similar for all groups. The median values increased from the incremental group (5.0) to the control group (6.0) and the DDA group (6.0). These differences were not statistically significantly different, $\chi 2(3) = 3.628$, $p = .163$.

We also considered relationships between flow, anxiety, challenge, enjoyment, and overall mean score across all three groups. Flow was positively correlated with Anxiety ($r(221) = .462$, $p < .0005$), Challenge ($r(221) = .476$, $p < .0005$), and Enjoyment ($r(221) = .675$, $p < .0005$), Anxiety was positively correlated with Challenge ($r(221) = .531$, $p < .0005$) and Enjoyment ($r(221) = .511$, $p < .0005$), and Challenge was positively correlated with Enjoyment ($r(221) = .501$, $p < .0005$). Overall mean score was significantly negatively correlated with Challenge ($r(221) = -.181$, $p = .007$).

In the case of hypothesis H1 – that DDA will produce greater player performance than either incremental or no difficulty adjustment – we are able to reject the null hypothesis. However, for H2 and H3, we are unable to reject the null hypotheses. That is, we cannot reject the hypotheses that there is no difference between DDA, incremental difficulty adjustment, and the control group in terms of either player experience of flow or player enjoyment.

Discussion and Conclusion

The research presented here investigated dynamic difficulty adjustment in a digital game, with the adjustment based on feedback about the player's performance. We found that this type of difficulty adjustment led to greater overall player performance over approximately 15 minutes of play, than either incremental difficulty adjustment or a no-adjustment control group. In addition, the range of performance in the DDA group was smaller than the other groups, and overall performance in the DDA group correlated with overall mean difficulty across 15 minutes of play. In line with previous research, our results suggest that performance-based DDA is suitable for reducing skill differentials between players. This provides further evidence of the suitability of this relatively simple approach to DDA in facilitating competitive and or collaborative play between players with different skill levels. We believe that this approach could therefore be used to increase the accessibility of digital games for people with disabilities.

It is also interesting to note that the players in the DDA group showed a wide range of abilities, as indexed by the range of difficulty settings recorded during the experiment (Figure 4). However, the overall mean performance of this group increased more than both the control and incremental groups, with significant differences in mean performance in most levels of the game between the DDA and other groups. It is therefore possible, in line with the results of Sampayo-Vargas et al.[50] that DDA provides a scaffold to players of a range of abilities, by making the game easier when their performance drops; this scaffold is then removed when their performance increases.

Our results show no significant differences between the three conditions on all the self-reported measures of player experience (flow, enjoyment, challenge, and anxiety). Previous research on performance-based DDA has found mixed results on self-report measures of player experience such as enjoyment, engagement, motivation, and challenge. There are several possible explanations for this. Firstly, it may be the case that performance-based approaches to DDA have less or no effect on self-reported player experience than affective approaches. This is feasible, as several studies have found that player perceptions of gameplay experience are related to factors other than player skill or performance, such as whether the users are experienced or casual players of digital

50 SAMPAYO-VARGAS, S. et al.: The Effectiveness of Adaptive Difficulty Adjustments on Students' Motivation and Learning in an Educational Computer Game. In *Computers & Education*, 2013, Vol. 69, No. 10, p. 459-460.

games,[51] their motivations to play games, their personality, or gameplay preferences.[52] If the aim of DDA is to increase players' positive perception of the gameplay experience, then affective approaches to DDA may be more useful. Secondly, within performance-based DDA approaches, there is a large range of factors which could influence player experience. In this study, we used simple measures of player performance to alter a single variable affecting game difficulty. There are many other factors we could have chosen. In addition, we could have provided a different range of difficulty settings (e.g., with larger or smaller increments), used a different target success rate, adapted the difficulty more or less frequently, and so on. These adjustments could lead to different results, and future research should consider, not just the difference between adaptive and non-adaptive difficulty adjustment, but also differences between alternative approaches to DDA. Thirdly, as discussed in the Methods section, self-report data obtained from MTurk may be less reliable than data obtained from traditional sources. While we included attention-check items to identify participants who were potentially selecting random answers to the questions, found high reliability for our questionnaire, and found expected correlations between flow, anxiety, challenge, and enjoyment, we did not use any other techniques to identify participants who were not engaged with the questionnaire. For example, it has been suggested that MTurk data can be made more reliable by explicitly asking participants if they answered the questions genuinely and assuring them that they will still be paid if they admit they did not.[53] Note that this limitation is somewhat mitigated in this study as we removed responses from participants whose performance data indicated that they had not engaged with the game. However, it is still feasible that some participants engaged with the game and read the questions but still provided unreliable data simply by not providing considered answers.

This study makes several contributions to research on dynamic difficulty adjustment in digital games. Our results show that performance-based dynamic difficulty adjustment can be used to increase player performance and reduce performance differentials in a 2D digital game. We also demonstrate the feasibility and limitations of conducting digital games research using an experimental game via Amazon Mechanical Turk. We make recommendations for future research to further investigate the effects of dynamic difficulty adjustment on enjoyment and experience of flow (for which we found non-significant results), such as using different feedback measures (including affective feedback), adjusting different game variables, and implementing additional steps to ensure the reliability of data gathered by player self-report.

BIBLIOGRAPHY

ADAMS, E.: *The Designer's Notebook: Difficulty Modes and Dynamic Difficulty Adjustment*. Released on 14th May 2008. [online]. [2022-05-13]. Available at: <https://www.gamasutra.com/view/feature/132061/the_designers_notebook_.php>.

ALEXANDER, J. T., SEAR, J., OIKONOMOU, A.: An Investigation of the Effects of Game Difficulty on Player Enjoyment. In *Entertainment Computing*, 2013, Vol. 4, No. 1, p. 53-62. ISSN 1875-9521.

ALTIMIRA, D. et al.: Enhancing Player Engagement Through Game Balancing in Digitally Augmented Physical Games. In *International Journal of Human-Computer Studies*, 2017, Vol. 103, No. 7, p. 35-47. ISSN 1071-5819.

51 ALEXANDER, J. T., SEAR, J., OIKONOMOU, A.: An Investigation of the Effects of Game Difficulty on Player Enjoyment. In *Entertainment Computing*, 2013, Vol. 4, No. 1, p. 61.

52 KARPINSKYJ, S., ZAMBETTA, F., CAVEDON, L.: Video Game Personalisation Techniques: A Comprehensive Survey. In *Entertainment Computing*, 2014, Vol. 5, No. 4, p. 212-213.

53 ROUSE, S. V.: A Reliability Analysis of Mechanical Turk Data. In *Computers in Human Behavior*, 2015, Vol. 43, No. 2, p. 306-307.

ANG, D., MITCHELL, A.: Comparing Effects of Dynamic Difficulty Adjustment Systems on Video Game Experience. In SCHOUTEN, B., MARKOPOULOS, P., TOUPS, Z. (eds.): *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. New York : ACM, 2017, p. 317-327.
ANG, D., MITCHELL, A.: Representation and Frequency of Player Choice in Player-Oriented Dynamic Difficulty Adjustment Systems. In ARNEDO, J., NACKE, L. E. (eds.): *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. New York : ACM, 2019, p. 589-600.
BARRETT, N. et al.: The Use and Effect of Video Game Design Theory in the Creation of Game-Based Systems for Upper Limb Stroke Rehabilitation. In *Journal of Rehabilitation and Assistive Technologies Engineering*, 2016, Vol. 3, No. 1, p. 1-16. ISSN 2055-6683. [online]. [2022-05-13]. Available at: <https://journals.sagepub.com/doi/full/10.1177/2055668316643644>.
BATEMAN, S. et al.: Target Assistance for Subtly Balancing Competitive Play. In TAN, D. (ed.): *Proceedings of the SIGCHI Conference on Human Factors in Computing System*. New York : ACM, 2011, p. 2355-2364.
BEVILACQUA, F., ENGSTRÖM, H., BACKLUND, P.: Game-Calibrated and User-Tailored Remote Detection of Stress and Boredom in Games. In *Sensors*, 2019, Vol. 19, No. 13, p. 1-43. ISSN 1424-8220. [online]. [2022-05-13]. Available at: <https://www.mdpi.com/1424-8220/19/13/2877/htm>.
BONTCHEV, B.: Adaptation in Affective Video Games: A Literature Review. In *Cybernetics and Information Technologies*, 2016, Vol. 16, No. 3, p. 3-34. ISSN 1311-9702.
CASLER, K., BICKEL, L., HACKETT, E.: Separate but Equal? A Comparison of Participants and Data Gathered via Amazon's MTurk, Social Media, and Face-to-Face Behavioral Testing. In *Computers in Human Behavior*, 2013, Vol, 29, No. 6, p. 2156-2160. ISSN 0747-5632.
CRUMP, M. J., McDONNELL, J. V., GURECKIS, T. M.: Evaluating Amazon's Mechanical Turk as a Tool for Experimental Behavioral Research. In *PLOS One*, 2013, Vol. 8, No. 3, p. 1-18. ISSN 1932-6203. [online]. [2022-05-13]. Available at: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0057410>.
CSIKSZENTMIHALYI, M.: The Flow Experience and Its Significance for Human Psychology. In CSIKSZENTMIHALYI, M., CSIKSZENTMIHALYI, I. S. (eds.): *Optimal Experience: Psychological Studies of Flow in Consciousness*. Cambridge : Cambridge University Press, 1988, p. 15-35.
ENGESER, S., RHEINBERG, F.: Flow, Performance and Moderators of Challenge-Skill Balance. In *Motivation and Emotion*, 2008, Vol. 32, No. 3, p. 158-172. ISSN 0146-7239.
FLEISCHER, A., MEAD, A. D., HUANG, J.: Inattentive Responding in MTurk and Other Online Samples. In *Industrial and Organizational Psychology*, 2015, Vol. 8, No. 2, p. 196-202. ISSN 1754-9426.
GERLING, K. M. et al.: Effects of Balancing for Physical Abilities on Player Performance, Experience and Self-Esteem in Exergames. In JONES, M., PALANQUE, P. (eds.): *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*. New York : ACM, 2014, p. 2201-2210.
HERNANDEZ, H. A. et al.: Designing Action-Based Exergames for Children with Cerebral Palsy. In MACKAY, W. (ed.): *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York : ACM, 2013, p. 1261-1270.
HUNICKE, R.: The Case for Dynamic Difficulty Adjustment in Games. In LEE, N. (ed.): *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology*. New York : ACM, 2005, p. 429-433.
HWANG, S. et al.: How Game Balancing Affects Play: Player Adaptation in an Exergame for Children with Cerebral Palsy. In MIVAL, O. (ed.): *Proceedings of the 2017 Conference on Designing Interactive Systems*. New York : ACM, 2017, p. 699-710.
CHEN, J.: Flow in Games (And Everything Else). In *Communications of the ACM*, 2007, Vol. 50, No. 4, p. 31-34. ISSN 0001-0782.
CHEUNG, J. H. et al.: Amazon Mechanical Turk in Organizational Psychology: An Evaluation and Practical Recommendations. In *Journal of Business and Psychology,* 2017, Vol. 32, No. 4, p. 347-361. ISSN 0889-3268.

JACKSON, S. A., MARSH, H.: Development and Validation of a Scale to Measure Optimal Experience: The Flow State Scale. In *Journal of Sport & Exercise Psychology*, 1996, Vol. 18, No. 1, p. 17-35. ISSN 0895-2779.

KARPINSKYJ, S., ZAMBETTA, F., CAVEDON, L.: Video Game Personalisation Techniques: A Comprehensive Survey. In *Entertainment Computing*, 2014, Vol. 5, No. 4, p. 211-218. ISSN 1875-9521.

LEIKER, A. M., et al.: The Effects of Autonomous Difficulty Selection on Engagement, Motivation, and Learning in a Motion-Controlled Video Game Task. In *Human Movement Science*, 2016, Vol. 49, No. 5, p. 326-335. ISSN 0167-9457.

NAGLE, A. et al.: The Effect of Different Difficulty Adaptation Strategies on Enjoyment and Performance in a Serious Game for Memory Training. In CLUA, E., VILAÇA, J. (eds.): *IEEE 3rd International Conference on Serious Games and Applications for Health*. Rio de Janeiro : IEEE, 2014, p. 127-134.

ORVIS, K. A., HORN, D. B., BELANICH, J.: The Roles of Task Difficulty and Prior Videogame Experience on Performance and Motivation in Instructional Videogames. In *Computers in Human Behavior*, 2008, Vol. 24, No. 5, p. 2415-2433. ISSN 0747-5632.

PERTTULA, A. et al.: Flow Experience in Game Based Learning – A Systematic Literature Review. In *International Journal of Serious Games*, 2017, Vol. 4, No. 1, p. 57-72. ISSN 2384-8766.

PIERRE-LOUIS, S.: *2021 Essential Facts About the Video Game Industry*. 2021. [online]. [2022-05-13]. Available at: <https://www.theesa.com/wp-content/uploads/2021/08/2021-Essential-Facts-About-the-Video-Game-Industry-1.pdf>.

RHEINBERG, F.: *Flow Short Scale*. [online]. [2022-05-13]. Available at: <http://www.psych.uni-potsdam.de/people/rheinberg/messverfahren/fks1-e.html>.

RHEINBERG, F., VOLLMEYER, R., ENGESER, S.: Die Erfassung des Flow-Erlebens. In STIENSMEIER-PELSTER, J., RHEINBERG, F. (eds.): *Diagnostik von Motivation und Selbstkonzept: Tests und Trends der pädagogisch-psychologischen Diagnostik – Band 2.* Göttingen : Hogrefe, 2003, p. 261-279.

ROBB, N., WALLER, A., WOODCOCK, K. A.: Developing a Task Switching Training Game for Children with a Rare Genetic Syndrome Linked to Intellectual Disability. In *Simulation & Gaming*, 2019, Vol. 50, No. 2, p. 160-179. ISSN 1046-8781.

ROUSE, S. V.: A Reliability Analysis of Mechanical Turk Data. In *Computers in Human Behavior*, 2015, Vol. 43, No. 2, p. 304-307. ISSN 0747-5632.

SAMPAYO-VARGAS, S. et al..: The Effectiveness of Adaptive Difficulty Adjustments on Students' Motivation and Learning in an Educational Computer Game. In *Computers & Education*, 2013, Vol. 69, No. 10, p. 452-462. ISSN 0360-1315.

SHAKER, N., YANNAKAKIS, G., TOGELIUS, J.: Towards Automatic Personalized Content Generation for Platform Games. In YOUNGBLOOD, G. M., BULITKO, V. (eds.): *Sixth Artificial Intelligence and Interactive Digital Entertainment Conference*. Burnaby : PKP Publishing Services, 2010, p. 63-68. [online]. [2022-05-13]. Available at: <https://ojs.aaai.org/index.php/AIIDE/article/view/12399>.

SHAREK, D., WIEBE, E.: Investigating Real-time Predictors of Engagement: Implications for Adaptive Videogames and Online Training. In *International Journal of Gaming and Computer-Mediated Simulations*, 2015, Vol. 7, No. 1, p. 20-37. ISSN 1942-3888.

SHERRY, J. L.: Flow and Media Enjoyment. In *Communication Theory*, 2004, Vol. 14, No. 4, p. 328-347. ISSN 1050-3293.

THIBAULT, R. T., LIFSHITZ, M., RAZ, A.: The Self-Regulating Brain and Neurofeedback: Experimental Science and Clinical Promise. In *Cortex*, 2016, Vol. 74, No. 1, p. 247-261. ISSN 0010-9452.

XUE, S. et al.: Dynamic Difficulty Adjustment for Maximized Engagement in Digital Games. In BARRET, R., CUMMINGS, R. (eds.): *WWW '17 Companion: Proceedings of the 26th International Conference on World Wide Web Companion*. Geneva : International World Wide Web Conferences Steering Committee, 2017, p. 465-471.

ZOHAIB, M.: Dynamic Difficulty Adjustment (DDA) in Computer Games: A Review. In *Advances in Human-Computer Interaction*, 2018, Vol. 11, No. 1, p. 1-12. ISSN 1687-5907. [online]. [2022-05-13]. Available at: <https://www.hindawi.com/journals/ahci/2018/5681652/>.