ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/357025764

Systematic Review of Dynamic Difficulty Adaption for Serious Games: The Importance of Diverse Approaches

Article *in* SSRN Electronic Journal · January 2021

DOI: 10.2139/ssrn.3982971

CITATIONS
13

READS
922

2 authors:

Andrew Seyderhelm
University of Newcastle Australia
6 PUBLICATIONS 33 CITATIONS

SEE PROFILE

Karen Blackmore
University of Newcastle Australia
78 PUBLICATIONS 1,339 CITATIONS

SEE PROFILE

All content following this page was uploaded by Andrew Seyderhelm on 09 June 2022.

The user has requested enhancement of the downloaded file.

Systematic Review of Dynamic Difficulty Adaption for Serious Games: The Importance of Diverse Approaches.

Andrew J A Seyderhelm[0000-0003-0191-0367] and Karen L Blackmore[0000-0002-9111-0293]

School of Electrical Engineering & Computing, University of Newcastle, Callaghan, NSW Australia

Andrew.Seyderhelm@uon.edu.au

Abstract. This study presents a systematic review of the literature on dynamic difficulty adjustment (DDA) used in games, in the context of serious games and simulation training. We consider factors such as the types of adaption strategies, the measures used to trigger adaptions, and the reported effectiveness of the approach to deliver improved performance and engagement. Five key bibliometric databases were surveyed and a total of 59 research articles were selected for analysis. From the analysis of these articles, it was identified that 36% included some form of physiological measurement. Approaches that adapted only a single component of the game were the dominant design, with 37% of studies analysed adopting multiple adaption strategies. Of particular interest for serious games are approaches that assist in providing a mechanism for dynamic difficulty adjustment based on cognitive performance. Results from the analysis indicate that a dynamic difficulty adjustment approach that incorporates multiple measures of performance are likely to result in improved learning outcomes. These findings contribute to existing literature on the design of games for learning and training and provide important directions for future work.

Keywords: affective computing, dynamic difficulty adjustment, serious games, adaptive, simulations, systematic literature review.

Declarations:

Funding: This work was supported by the Defence Innovation Network NSW PhD Grant, and the University of Newcastle Postgraduate Research Scholarship.
Conflicts of interest/Competing interests: Not Applicable
Availability of data and material: Not Applicable
Code availability: Not Applicable
Authors' contributions: All authors contributed to the study and design. Literature search and data analysis performed by Andrew J A Seyderhelm. Review of results (PRISMA) performed by Andrew J A Seyderhelm and Karen L Blackmore. Writing – original draft: Andrew J A Seyderhelm; Writing – review and editing: Andrew J A Seyderhelm and Karen L Blackmore; Critical review & Supervision: Karen L Blackmore.

1 Introduction

Numerous studies have shown that serious games are broadly more effective than conventional instruction (Connolly, Boyle, MacArthur, Hainey, & Boyle, 2012; Sitzmann, 2011; Wouters, van Nimwegen, van Oostendorp, & van der Spek, 2013). Adaptivity, or the ability for a game to change in response to the player, in serious games has been beneficial and helps promote "individualized learning" (Ravyse, Seugnet Blignaut, Leendertz, & Woolner, 2016), however only a relatively small proportion of serious games employ adaptivity (Ravyse, et al., 2016).

Understanding effective adaptive mechanisms and approaches is important to improve training and learning outcomes. Adaptions can be made based on prior understanding of the player, however, this does not account for progress and development in real-time, and thus dynamic approaches are of particular interest. In this work, we present a systematic review of relevant literature to assess the current state of Dynamic Difficulty Adjustment (DDA) and adaptive approaches in games. In doing this, we aim to create a snapshot of the current state of the art and seek to assess which of the existing techniques may be best applied to serious games and simulation training.

Serious games is a common umbrella term used to describe any game with a serious purpose where the primary goal is not entertainment; this purpose is usually some form of education or training (Wouters, et al., 2013). Additionally, Kapp (2012) defines a serious game as "[…] an experience designed using game mechanics and game thinking to educate individuals in a specific content domain" (p.15). Conversely, simulation training refers to an "[…] artificial representation of a real-world process to achieve educational goals through experiential learning" Al-Elq (2010) (p.35). However, both simulations and serious games may benefit from dynamic difficulty adjustments.

Historically, the primary purpose of DDA in entertainment gaming is to create a more enjoyable experience by matching the player to the optimal difficulty or challenge (Hunicke & Chapman, 2004; Xue, Wu, Kolen, Aghdaie, & Zaman, 2017). Traditionally, video games provide a discrete number of difficulty levels that the player will select before playing; this could range from easy to extremely hard and the difficulty level in a game may change over time based on game design considerations. Player performance can change over time based on a plethora of factors, DDA systems can recognise these changes and adapt to the evolving player needs. Therefore, the aim of DDA is to help players achieve an optimal state of challenge and thus maintain interest.

The aim of serious games may be best achieved by engaging learners actively in the learning material (Ravyse, et al., 2016). Hamari, et al. (2016) suggests serious games can enhance learning experiences through appropriate levels of challenge and skill. This latter point aligns to DDA concepts from entertainment games, but there is need to extend our understanding of the value of DDA in serious games, incorporating the notion that matching appropriate skill to challenge is important for maximising the benefits of serious games and accounting for learning or training goals. Furthermore, adapting the challenge of the game commensurate with increasing player skill is correlated with both improved engagement and learning outcomes (Hamari, et al., 2016).

Real-time adaption in a learning-based setting has the potential to provide more effective training that adapts to the needs of the students and more closely approximates one-on-one training (Landsberg, Van Buskirk, Astwood Jr, Mercado, & Aakre, 2010). From this perspective, there is potential for advances in DDA, from the entertainment gaming space, to provide insights into methods or approaches that can increase learning and training outcomes in serious gaming and simulation contexts. We first begin by providing an overview of existing research and reviews in this area prior to outlining the methodology adopted for our review. Following this, we present the findings of the systematic review, and outline their applicability to serious games and simulation training using a thematic analysis framework. From this, we draw conclusions and outline areas for future work.

2 Previous research

This systematic review focusses on identifying different approaches to the implementation of dynamic difficulty adjustment (DDA) in games, and the application of these approaches to serious games and simulation training. Two key areas of previous research provide the framework for this review. Firstly, existing reviews of DDA are important sources for identifying potential gaps in coverage and scope that would be useful to address. Secondly, given the focus on enhancing learning outcomes through game-based approaches, a lens for analysing the results of this review that captures this intent, is also needed. In this context, learning reflects concepts such as competence and mastery, and thus the utility of DDA approaches to facilitate key aspects from relevant learning theories is also considered. The following review sections address these specific areas.

2.1 Existing dynamic difficulty adjustment reviews

DDA and adaptive techniques have been topics of significant research over the last decade. This interest has principally focussed on entertainment gaming, although adaptive techniques have also been used in serious games. Ravyse, et al. (2016) conducted a systematic review producing guidance on how to make effective serious games; they developed five themes for successful games, one of which included adaptivity. Additionally, they found that only 21% of the papers they reviewed included adaptive mechanisms, an amount they refer to as a “disappointingly small proportion” (Ravyse et al, 2016. p. 50), outlining a gap in the existing body of research. Additionally, while focussed on serious games, the review does not provide a holistic review of dynamic difficulty adjustment approaches.

Two additional literature reviews deal with adaptivity in games (Bontchev, 2016; Zohaib, 2018). Zohaib (2018) exclusively focusses on DDA methods, considering papers on DDA from 2009 to 2018. However, this review did not specify any process or criteria, and thus a clear research question is not obvious. The review considered 84 papers between the afore-mentioned dates representing a sound overview of the state of DDA. Many of the reviewed articles were conceptual and did not validate the proposed DDA methods through human-based testing. Zohaib provides a good summary of the various DDA methods reviewed, but a discussion of the effectiveness or usefulness of the methods for different purposes is lacking. This is confirmed by Zohaib, stating that "[T] the number of approaches presented here is neither complete nor exhaustive but merely a sample that demonstrates the usefulness and possible applications of AI techniques in modern video games. (p. 10)"

Bontchev (2016) literature review focusses on affective adaption methods and does not cover the full scope of DDA methods. However, Bontchev's paper does provide an overview of DDA methods and a more detailed description of affective adaptive techniques from the 14 games included in the review. Bontchev's review also included analysis of each paper's efficacy and links to psychological theorem. A detailed overview of various measuring techniques, how well they worked and how intrusive they were, was also provided. Finally, Bontchev provided a clear summary and recommendations for future research direction. Irrespective of its thorough analysis, this review was of limited scope due to the sole focus on affective adaption methods, and thus doesn't inform broader questions relating to alternative approaches for DDA in learning and educational game contexts.

Finally, in addition to the two papers outlined above, most articles reviewed in this paper devote a portion of their writing to an overview of the literature. However these literature review sections are frequently narrow and pertain predominantly to the specificity of the contained research. Thus, the review presented here provides a novel assessment of DDA approaches in the context of their use in serious games and simulation training, and addresses a stated gap in knowledge in this area (Ravyse et al, 2016).

2.2 Relevant learning and psychological theories

There are relevant frameworks that can be used in the implementation of adaption mechanisms to achieve improved learning outcomes from a learning theory and/or psychological perspective. Self-determination theory (SDT) outlines the importance of an individual's competence, autonomy and relatedness in achieving improved motivation and healthy psychology (Richard M Ryan & Deci, 2017). It can relate to learning in games as it defines how we seek to gain mastery of challenges to achieve a sense of competence which in turn improves motivation. Mastery of a game can lead to greater enjoyment (Richard M. Ryan, Rigby, & Przybylski, 2006) and similarly, greater engagement. Negini, Mandryk, and Stanley (2014) showed how using DDA methodology to give players the support to master challenges, as opposed to making the challenges intrinsically easier, is more effective as it maintains a sense of autonomy while supporting the achievement of competence; this correlates to SDT (Richard M Ryan & Deci, 2017). Additionally, this aligns with Vygotsky's Zone of Proximal Development (ZPD) (Kozulin, Ageyev, Gindis, & Miller, 2003). ZPD is often simplified to refer to the difference between what a student can achieve alone versus what can be achieved when guided by a more knowledgeable or experienced person (Kozulin, et al., 2003). In serious games and simulation training, this guidance may be fulfilled by a non-player character (Lim, et al., 2014). Competence is easily derailed by challenges that are too difficult, or those that provide excessive negative feedback (Richard M Ryan & Deci, 2017). Therefore, a DDA system that supports the individual and adjusts to attain the ideal challenge should help engender mastery or competence. This parallels Csikszentmihalyi, Abuhamdeh, and Nakamura (2014) theory of flow in which the right level of challenge must be met in order to achieve a heightened state of focus and engagement (flow), thus avoiding boredom or frustration.

Another approach is described by Cognitive Load Theory (CLT), where expertise and competence are directly correlated to long-term memory and schema development; the greater the expertise the more developed the schema (Sweller, 2011). From this perspective, once a topic or task has become automatic, it is said to be stored in long-term memory as schema and then correlates with a drop in cognitive workload to undertake that task (Paas, Renkl, & Sweller, 2004). Cognitive load is an ideal measure to combine with performance metrics, as it details how hard a person's cognitive processes are working to achieve the measured level of performance. This is critical to understand in developing adaptive serious games and simulation training (Dideriksen, et al., 2018), because as a participant achieves competency at a certain challenge level, the cognitive resources needed to achieve the required performance will reduce. In other words, performance should go up as cognitive load reduces (Aldekhyl, Cavalcanti, & Naismith, 2018; Sweller, 2011). By measuring cognitive load and performance, the ideal balance can be struck to maintain the optimal level of challenge and engagement. If a participant is achieving the required performance level, and they do this with a low cognitive load, they can be said to have achieved the desired level of mastery of the subject matter; this has been observed in medical simulation training (Haji, Rojas, Childs, de Ribaupierre, & Dubrowski, 2015) and is a core tenet of CLT (Sweller, 2011). Additionally, it has been hypothesised that for a learner to transition from a training environment to a less controlled, or more complex, real-world environment, they need to achieve the performance targets with a low cognitive load (Aldekhyl, et al., 2018; Dideriksen, et al., 2018). CLT therefore highlights the importance of recognising cognitive performance associated with tasks or activities in the learning process, and thus has the potential, when combined with performance, to inform the adaption process.

2.3 DDA in education and training environments

When seeking to create a DDA system for serious games and simulation training, it is imperative to understand the structure of the game, how it fits into the existing curriculum, and how best to reflect training and learning principles (Seyderhelm, Blackmore, & Nesbitt, 2019). The result should be a well-integrated, flexible and adaptable game that closely approximates one-on-one training and is tightly integrated into the curriculum (Popescu, Roceanu, Earp, Ott, & Ger, 2012).

It is well established that serious games and simulation training are effective (Connolly, et al., 2012; De Freitas, 2018) when appropriately integrated into curriculum (Wouters, et al., 2013). Sitzmann (2011) highlighted that a key to educational success is "[…] maximised when trainees actively rather than passively learned work related-competencies during game play" (p.520), which describes the purpose of DDA. Sitzmann (2011) goes on to suggest "[…] additional research is needed to examine the dynamic interplay of affective and cognitive processes during game play and, ultimately, their effect on training transfer." (p.520). Finally, Vlachopoulos and Makri (2017) suggest further research should focus on "[…] applying relevant theoretical frameworks […] to cognitive, behavioural and affective outcomes, respectively" (p. 27). This is a call taken up in this literature review and explored further in terms of DDA methodologies and applications. It is from this perspective that the review presented here seeks to provide a more comprehensive and targeted overview of existing approaches to DDA. In the following section, we outline the methodological approach adopted for the review, including the research questions that frame the resultant analysis.

3 Method

PRISMA guidelines were used as the basis of this systematic literature review (Moher, et al., 2015), with the following subsections reflecting the process stages. An initial scoping search of existing research revealed many theoretical and untested-on-human concepts and approaches. From this scoping search, the key inclusion criteria for this literature review was defined. Of most importance is that the included DDA systems are validated on real people and therefore provide experiment results. Also, given the fundamental differences between real-time and a-priori adaption mechanisms, and the relevance of real-time adaption in learning contexts (Kelley, 1969), exclusion criteria were also framed around this concept. Review criteria were therefore restricted to include only studies where the adaption occurs in real-time and, to ensure that the broadest possible scope of papers relevant to the research was captured, the search was not date limited.

3.1 Research question

As per the PRISMA guidelines (Moher, et al., 2015), the systematic review is guided by an overarching research question:

> *What approaches for dynamic difficulty adjustment (DDA) exist and might they apply to serious games and simulation training?*

We expect DDA for entertainment games to adapt the game mechanics and environment, whereas DDA for serious games may focus more on learning content. It is critical to understand these differences when assessing the most appropriate methods of DDA for serious games and simulations. At one end of the spectrum, a relatively simple 2D game teaching student's basic knowledge of bacteria and lab tests (De la Cruz, Portocarrero, & Shiguihara-Juárez, 2018) may adapt the questions and quizzes in the game based on player performance rather than the gameplay elements. Whereas at the other end of the spectrum, player inventory, health, damage, accuracy, enemy hit points and more (Hunicke, 2005) can adapt a game to make it more enjoyable for players. There is a large overlap in this spectrum, describing the zone where both the learning content and the game-play content is adapted (Fig. 1). The research question seeks to explore this overlap space more fully.

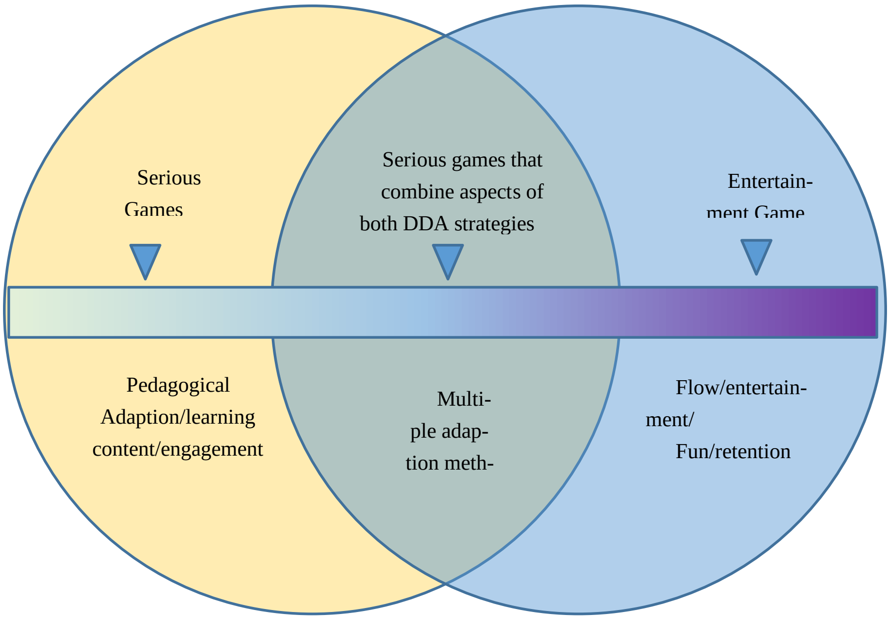

Fig. 1. Serious game to entertainment game adaption continuum.

3.2 Information sources and search

This study involved searches in bibliographic databases pertinent to computer science, entertainment gaming, education, and information technology (Table 1). Detailed searches were conducted in the databases listed in Table 1 on the 21st and 22nd of February 2019 and updated on 28 October 2020 and again on 20 January 2021, and included conference papers, book chapters and journal articles. The following terms were used for the searches (in various combinations dependent on the database search functions), resulting in a total of 1916 results: "dynamic difficulty", "adapt* game", "real-time difficulty adjust*", "automatic difficulty adjust*", "automatic difficulty balanc*", "adapt* difficulty", "serious game", "simulation", "game", "virtual", "computer","balanc*".

Table 1. Results of the search in five databases.

Database	# Results
World of Science	292
SCOPUS	722

IEEE	514
ACM Digital Library	158
The ACM Guide to Computing Literature (2019)	230
Total	**1916**

3.3 Study selection & data collection processes

Initial scanning of the 1916 search results revealed many articles related to working with learning difficulties, physical rehabilitation or disabilities, which included terms like “adapting games”; this is not adaptive gaming in the context of DDA as it refers to accessibility, for example, to be usable by a person with a disability or injury. Similarly, after reviewing many papers specific to physiotherapy and related practices, it was determined that adaption mechanisms in this field were often not applicable to the broader serious games and simulation training context, a point made by (Andrade, Pasqual, Caurin, & Crocomo, 2016):

> “It is important to note that games developed for rehabilitation differ from commercial entertainment games due to severe limitations imposed to patients by pathologies such as strokes, cerebral palsy and spinal cord injuries[.]”

As such, physiotherapy specific adaption using custom control schema were excluded from the study. This process of initial scanning was used to develop a set of exclusion terms (Table 2), which were applied to the returned search results (1916 articles) obtained rather than limiting the initial bibliographic database search. Careful exclusion of papers based on these terms reduced the set of relevant search results to 373 articles.

Table 2. Exclusion terms.

Radio	Power
Energy	Autis*
Rehab*	Handicap*
Child*	Physiotherapy
Exergam*, exert*	Tutor*
Personality	Therap*
Stochastic	Teaching NOT dynamic difficulty
Alz*	802 NOT dynamic Difficulty
Patho*	

The final review of the 373 relevant studies (Figure 2) involved double blind screening of all 373 abstracts, reducing the potentially relevant articles from 373 to 173. A similar double-blind approach was then taken based on a full-text review of the 173 potentially relevant articles. Any conflicts over exclusion criteria in both stages of review were resolved via discussion; an arbiter was identified to deal with any situations where agreement was unattainable. This process further reduced the number of relevant papers to 59. The exclusion criteria for this reduction are also provided in Figure 2.

Fig. 2. Systematic literature review process.

4 Analysis

This section details the characteristics of the included studies, consisting of the years of publication and quality assessment of the publication outlets. Also detailed are characteristics of the individual studies described in the articles, including details on sample sizes and participant demographics, and game genre where the DDA mechanism is applied. This information provides context for the subsequent review of the studies contained in the located articles.

4.1 Year of publication of articles

The 59 papers included in this review had publication dates ranging from 2005 – 2021 (Figure 3).

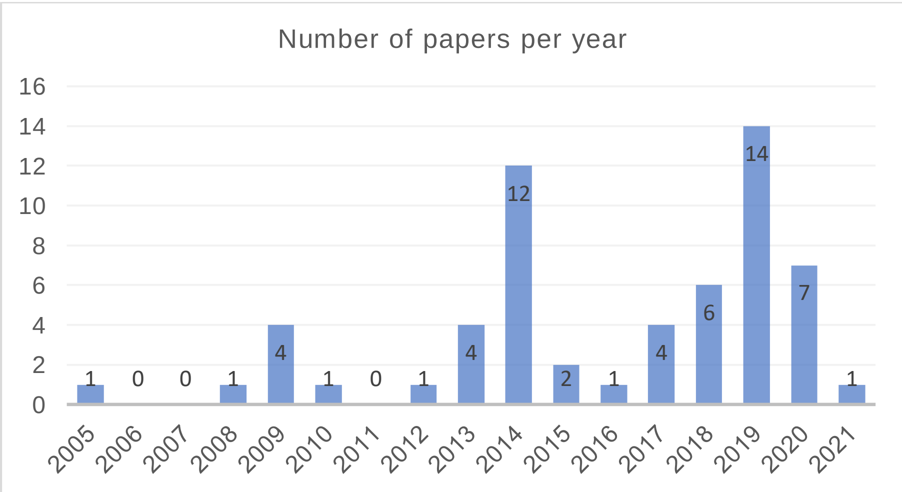

Fig. 3. Numbers and years of publishing.

4.2 Quality distribution of articles

There are several ranking systems available that can provide objective metrics on the quality of journal publication outlets (Gu & Blackmore, 2017). The Scimago Journal and Country Rank (SJR) score, developed from Scopus bibliometric data, provides a robust comparative measure of journal impact. Journals that have been ranked according to their SJR are divided into four equal groups, or quartiles (Q1, Q2, Q3, and Q4), with Q1 journals including the top 25% of journals in each field.

Twenty-one (35.6%) studies in this review were published in journals, with 38 (64.4%) articles published in books and/or conference proceedings. Journal publications can be assessed for impact and/or quality based on interquartile ranks, with the articles from this review distributed as follows: Q1 = seven (33%); Q2 = five (24%); Q3 = seven (33%); and Q4 = two (10%).

4.3 Participant sample sizes for experiments

Within the included articles, the number of experiment participants varied widely, with many studies having sample sizes below 30 (Figure 4).

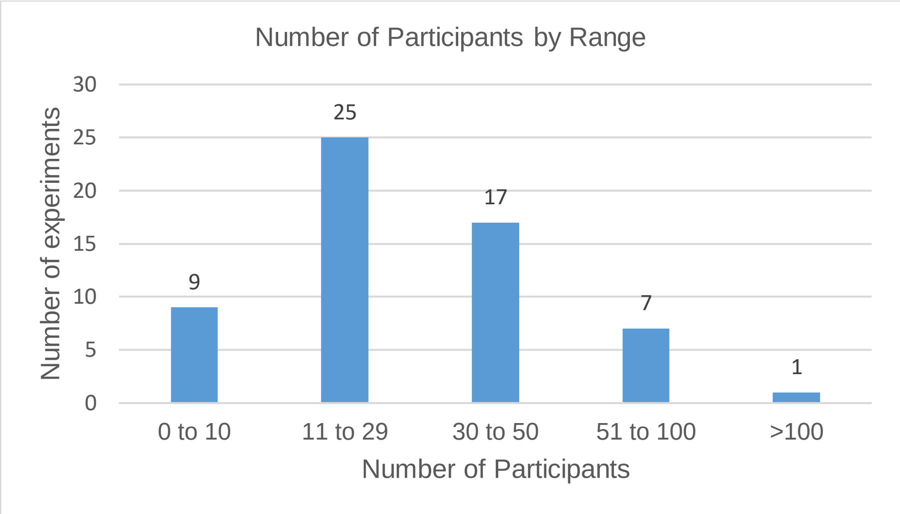

Fig. 4. Number of participants in experiments by range.

4.4 Participant Demographics

In total, 1878 participants were involved in the evaluations of DDA described in the articles reviewed. Across the 59 papers, 30 papers did not record the proportion of males and females in the experiment sample; of those that did, 62.9% were male. Twenty-six (25) papers did not record age information of participants, 24 recorded the mean age, and the remaining ten (10) reported an age range. Of those that reported a mean age, the overall result of articles reviewed was 26.27, six (6) reported an age range between 17-35, one 18-42, one 11-12, one 60-100 and one 6-65. 4.5 Game Genre and Type

The games featured in the selected articles consisted of two main categories of games: 2D (59.3%) and 3D (40.7%). The studies included 23 (39%) serious games, of which eleven (47.8%) were 2D games and the remaining twelve (52.2%) 3D. Conversely, the 36 (61%) remaining studies focussed on entertainment games, of which 24 (66.7%) were 2D games and 12 (33.3%) were 3D. The specific detail of each of the games, including their genre, are considered in Section 6 and Annexe A.

5 DDA & adaptive technique review

In the following sections we first discuss types of measures used in the included articles that form the basis for triggering a DDA mechanism (Section 5.1). We then summarise any noted educational or psychological theories that may underpin the described DDA systems. An analysis of the DDA adaption mechanisms was conducted leading to a list of nine (9) identified DDA mechanisms (Section 5.3).

5.1 Triggers for DDA Mechanisms

When considering DDA, there are two elements to each system. The first element relates to what is measured to trigger the DDA mechanism, whereas the second aspect is how the adaption mechanism is implemented. The articles included in our review describe games that adapt in real-time through a variety of measures and mechanisms. These can be broadly categorized as follows:

1. *Performance adaptation* – this relates to adaptation based on the performance of the player in the game. Examples included: amount of damage inflicted (Denisova & Cairns, 2015; Vicencio-Moreira, Mandryk, & Gutwin, 2014); player health (Baldwin, Johnson, & Wyeth, 2014; Hunicke, 2005; Suaza, Gamboa, & Trujillo, 2019); distance travelled while collecting objects (Alexander, Sear, & Oikonomou, 2013; Li, Yao, Li, & Zhang, 2014); driving speed and accuracy (Rietveld, Bakkes, & Roijers, 2014); correct answers (De la Cruz, et al., 2018; A Nagle, Novak, Wolf, & Riener, 2014; A. Nagle, Novak, Wolf, & Riener, 2015; Peirce, Conlan, & Wade, 2008; Sampayo-Vargas, Cope, He, & Byrne, 2013; van Oostendorp, van der Spek, & Linssen, 2014), understanding and player behaviour (Pfau, Smeddinck, & Malaka, 2020; Tang, Liang, Hare, & Wang, 2020), or levels and sections cleared (Masanobu, B, & Mikami, 2017).
2. *Physiological adaptation* – these systems used physiological systems to trigger the DDA mechanism. Examples of methods used include: EEG (Ewing, Fairclough, & Gilleade, 2016; Stein, Yotam, Puzis, Shani, & Taieb-Maimon, 2018; Tahai, Wallace, Eckhardt, & Pietroszek, 2019), functional near-infrared spectroscopy (Afergan, et al., 2014), eye tracking (Abdessalem & Frasson, 2017; Spiel, Bertel, & Kayali, 2019), heart rate (Ninaus, Tsarava, & Moeller, 2019), electrodermal activity (Chanel & Lopes, 2020; Liu, Agrawal, Sarkar, & Chen, 2009) and facial expression tracking (Bontchev & Georgieva, 2018; Burns & Tulip, 2017; Reidy, Chan, Nduka, & Gunes, 2020). In most cases, these approaches were used to imply a mental state, such as boredom, arousal, or excitement. Many of the studies used multiple methods to record physiological data (Bian, et al., 2019; Liu, et al., 2009).
3. *Cognitive adaptation* – includes systems that measure the workload of the brain, for example working memory (Afergan, et al., 2014); the adaptation is based on a concept of free mental processing capacity, or lack thereof.

In summary, 21 (35.6%) articles use physiological measurement systems, 43 (72.9%) studies measure some form of player performance metric, and one (1) study incorporated a measure of cognitive load (1.7%). Additionally, six of these reported articles (10.2%) used dual measures as the basis of adaption and are thus recorded in both sets of statistics, resulting in 65 implementations.

5.2 DDA & adaption by educational and psychological theorem

Of the 59 papers reviewed, 42 (71.2%) made a clear reference to either a psychological or educational theory. Forty (40) of these 42 papers (95%) reference flow theory (Csikszentmihalyi, et al., 2014), of which nine (9) refer to at least one additional theory, with only two articles referring to solely other theorem, incongruity theory (Van Lankveld, Spronck, Van Den Herik, & Rauterberg, 2009) and experiential learning theory (Bontchev & Georgieva, 2018). Of the forty papers that reference flow, twenty-four (24) are entertainment based, ten (10) are serious games, and six (6) are a subset of serious games specific to health practices. Thus 66.7% of entertainment games reference flow, 58.8% of non-health related serious games and 100% of health games reference flow. This demonstrates that flow theory (Csikszentmihalyi, et al., 2014) is the most predominant theory in DDA application research with 71.2% of all reviewed articles referencing flow.

The papers were further divided into three overarching categories to identify the underlying purpose for the DDA:

1. *Entertainment games (61% of results)* - any game that has entertainment as its primary purpose or research in which the DDA method is hypothesised to work for entertainment games.
2. *Serious games & simulation training (28.8% of results)* – any implementation with an educational or training basis, where the DDA or adaption system is targeting improved educational outcomes.
3. *Serious games for health (10.2% of results)* – these are adaption and DDA methods that may be applicable to a broader audience but are primarily designed to apply in health contexts, including mental health. The delineation between health and serious games has been created in order to match the definition by Kapp (2012), whereby serious games are used to "educate individuals", and health games focus on motivating a patient to undertake a certain activity (Alves, Gama, & Melo, 2018), improve health or physiological processes (El-Habr, Garcia, Paliyawan, & Thawonmas, 2019; Tadayon, Amresh, McDaniel, & Panchanathan, 2018) or enhance a decaying cognitive function, e.g. memory (A Nagle, et al., 2014; A. Nagle, et al., 2015; Reidy, et al., 2020).

5.3 Effect of DDA per type

In both 2D and 3D DDA, there are several strategies employed to adjust the difficulty of the activity that range from changing the environment (Rietveld, et al., 2014), adjusting Artificial Intelligence (AI) (Tan, Tan, Tan, & Tay, 2009), and providing player assistance (Vicencio-Moreira, et al., 2014). Nine (9) DDA mechanisms were characterized from the articles reviewed (Table 3).

Table 3. DDA type and description.

DDA Method	Description
Time adaption	A time factor is extended or decreased giving the player more or less to achieve an outcome.

Vary the task complexity	The principle mechanic or task is structured to provide easier and harder challenges, e.g. medical complexity of cases is made more or less complex (van Oostendorp, et al., 2014), Tetris blocks are picked to be easier or harder (Lora Ariza, Sánchez-Ruiz, & González-Calero, 2019) or driving speed is increased or decreased (Parnandi & Gutierrez-Osuna, 2014).
Procedural Content Generation (PCG) or modification	The design of the level is procedurally generated/adjusted in order to make the level easier or harder, e.g. adjusting platform distance and height in a side scrolling game (Alexander, et al., 2013) or top-down racing games varying track complexity (Rietveld, et al., 2014).
In-game Feedback	Feedback and hints are provided in-game to assist the player complete the task. This feedback can take many forms from sound effects to NPCs, or pedagogical agents (Peirce, et al., 2008).
Artificial Intelligence (AI)	The AI is adjusted to create easier or harder challenges, e.g. in a board game the AI can play harder or easier moves and strategies (Tan, et al., 2009) or an AI system to control the overall DDA strategy, including AI opponents (Hunicke, 2005).
Direct player assistance	This reduces or aids an element of the player action/performance to improve or limit outcomes. For example aiming assistance in a FPS (Vicencio-Moreira, et al., 2014) or player speed and access to weapons (Negini, et al., 2014).
Environment Content Adaption	This adjusts elements in the game which assist or hinder the player, not the level itself, but rather elements within the level. This may include adjusting the number of health pick-ups (Hunicke, 2005), or weather FX (Parnandi & Gutierrez-Osuna, 2014).
Speed	The speed of elements in the game, not the player, are increased or decreased, e.g. the speed of falling blocks in Tetris (Xiang, Yang, & Zhang, 2013)
Challenge Adjustment	The number of enemies, targets or obstacles is increased or decreased (Zook & Riedl, 2014) or the abilities of the enemy or obstacles (e.g. hit points or weapon damage) are increased or decreased (Yun, Shastri, Pavlidis, & Deng, 2009).

From the review, 16 (27.1%) articles described games that implement more than one type of DDA. Some of the DDA mechanisms overlap, particularly regarding the task complexity. In these cases, the guiding principle taken in this research was to use the author(s) primary stated aim as the category.

6 Discussion of DDA mechanisms

As previously mentioned, DDA implementations consist of both a trigger for when player activity or performance meets a certain threshold, and an adaptation mechanism that modifies the in-game experience of the player. In this section, we review the differing adaption mechanisms described in Table 3, with reference to serious games and simulation training.

6.1 Adapting Time

Time is potentially a powerful DDA technique for serious games and learning. By altering the amount of time a player has in-game, designers can significantly affect participant experience (Ravyse, et al., 2016). Galy, Cariou, and Mélan (2012) showed that time pressure, and task difficulty, increase mental workload; manipulating these elements may be an effective aspect of a well-rounded DDA mechanism for serious games. When combined with other DDA methods, this can help advanced students move on to more challenging tasks, potentially avoiding boredom. Conversely, increasing the time for students who are struggling may help to improve learning outcomes. However, time-based adaption approaches may impose external time constraints on the learning activity itself. For example, a game that slowed game progress to provide additional time for mastery may not be effective if the time allowed for the learning activity or class was not similarly extendable. This external time pressure may reduce performance (Conte, Scarsini, & Sürücü, 2016; Schwabe & Wolf, 2010) and increase extraneous cognitive load impacting learning as an unintended consequence (Galy, et al., 2012).

6.2 Varying the Task Complexity

Task complexity is defined in this research as the intrinsic and primary mechanic of the game, for example making a medical triage scenario more complex (van Oostendorp, et al., 2014), increasing or decreasing simulated UAVs in a flight path simulator (Afergan, et al., 2014), adjusting the speed and steering of a vehicle (Parnandi & Gutierrez-Osuna, 2014), puzzle complexity (Tahai, et al., 2019) choosing more complicated Tetris pieces (Lora Ariza, et al., 2019; Masanobu, et al., 2017). Task complexity was used by eighteen (18) games, eleven (11) of which used it as a sole adaption strategy. DDA mechanisms based on task complexity are designed uniquely for each serious game, where each game has a different learning or training goal, for example arithmetic operations (Mainieri, Azevedo, Braga, & Omar, 2018). This implies a DDA mechanism based on task complexity may be challenging to apply broadly, despite it being one of the more effective DDA implementations. Task complexity aligns closely with the technique of scaffolding in educational software. Reiser (2004) describes this "[A] as applied to software, scaffolding refers to cases in which the tool changes the task in some way so that learners can accomplish tasks that would otherwise be out of their reach." (p. 275). Several of the articles reviewed apply approaches beyond simple adjustment of game parameters (see Appendix A).

Parnandi and Gutierrez-Osuna (2014) emphasise that games that adapt based on physiological signals are "[…] not trivial" and have not gained traction in the game community; the reasons for this are diverse. Physiological measures often infer emotion that may not necessarily correlate with mastery or relate to successful learning outcomes. That is, how a person feels about a serious game or how engaged they are with it does not necessarily translate to learning or mastering of its content (van Oostendorp, et al., 2014). Additionally, these methods require extra equipment (Bian, et al., 2019; Bontchev & Georgieva, 2018; Liu, et al., 2009; Park, Sim, & Lee, 2014; Parnandi & Gutierrez-Osuna, 2014) or significant time for establishing a baseline (Liu, et al., 2009), making them potentially unsuited to classroom or training settings. Conversely, performance-based adaption mechanisms show a great deal of promise as demonstrated by serious game papers reviewed. Indeed, for serious game implementations where mastery is sought in a content domain, measuring performance is not only an ideal method for DDA, but an effective way to assess student proficiency (van Oostendorp, et al., 2014).

6.3 Procedural Content Generation Adaption

Four (4) of the six (6) procedural content generation DDA systems identified in the reviewed articles were based on 2D side scrolling games. Two of the successful procedural content generation systems were based on multiple measures of player performance (Alexander, et al., 2013; Rietveld, et al., 2014) with one successful based on a single physiological measure however advocating a dual measure for future work (El-Habr, et al., 2019). One dual measure system, time and distance travelled, was less successful (Hendrix, Bellamy-Wood, McKay, Bloom, & Dunwell, 2019) and systems using only one measure for adaption were also less successful (Burns & Tulip, 2017; Li, et al., 2014). All the games were 2D based games, suggesting that PCG may be unsuitable for more complex 3D environments. The success rating combined with comments by respective authors indicates that multiple measurement strategies may be preferable in serious games as they capture more dimensions of player activity, affect or performance.

6.4 Speed as an Adaption Strategy

Speed was used as a principal adaption strategy in five (5) papers (Andrew, Tjokrosetio, & Chowanda, 2020; Chanel & Lopes, 2020; Ewing, et al., 2016; Spiel, et al., 2019; Xiang, et al., 2013) with nine (9) others using it as one of several DDA techniques (Abdessalem & Frasson, 2017; Alexander, et al., 2013; Bian, et al., 2019; Bontchev & Georgieva, 2018; Burns & Tulip, 2017; Li, et al., 2014; Liu, et al., 2009; Plass, Homer, Pawar, Brenner, & MacNamara, 2019; Tadayon, et al., 2018). Speed was more effective in 2D games as opposed to more complex 3D games, possibly due to the simpler nature of the games, where the speed of the game is a major component of gameplay, such as in Tetris-style clones (Chanel & Lopes, 2020; Ewing, et al., 2016; Spiel, et al., 2019; Xiang, et al., 2013), or Flappy Bird clones (Li, et al., 2014). Speed is therefore a limited method for DDA in and of itself as not all games may have their difficulty effectively adjusted by speed alone, however it has shown to be effective as part of a broader suite of adaption methods with six successful multi-implementations out of nine (Abdessalem & Frasson, 2017; Alexander, et al., 2013; Bontchev & Georgieva, 2018; Liu, et al., 2009; Tadayon, et al., 2018) versus only two successes out of five for single methods (Andrew, et al., 2020; Xiang, et al., 2013). Specifically, speed may be effective for DDA in serious games and simulation training, when combined with other measures in situations such as driving simulations, where obstacles or targets may move at differing speeds (Bontchev & Georgieva, 2018).

6.5 Feedback as Adaption

Five (5) papers used feedback as an adaption technique (Hooshyar, et al., 2021; Papadimitriou, Chrysafiadi, & Virvou, 2019; Peirce, et al., 2008; Tahai, et al., 2019; Tang, et al., 2020), these five (5) games are all serious games. Three of the five (5) games used feedback as part of a number of adaption methods and were each successful, the remaining two utilised a single approach and were both unsuccessful. Two employed pedagogical agents to provide this feedback and direction aligning with the social (relatedness) component of Self-Determination Theory (SDT) (Kim, 2001). Richard M Ryan and Deci (2017) emphasise that the SDT concept of relatedness has been shown to extend to well-crafted non-player characters (NPCs), suggesting NPC's can help enhance the learning material. The remaining three (3) games provide feedback with visual and audio cues. Other social-based methods can be used in serious games to enhance learning experiences, including developing multiplayer gameplay, or synchronous classroom game activities, that encourage discussion of game experiences. Feedback in serious games can help meet SDT needs through DDA by incorporating well-crafted pedagogical agents. This also aligns to the ZPD concept of tutor or mentor guidance, and an adaptive pedagogical agent may be able to perform this role guiding students to success (Papadimitriou, et al., 2019; Peirce, et al., 2008).

6.6 Artificial Intelligence Adaption

Seven (7) articles used an AI opponent approach as the sole adaption mechanism, with one using it in conjunction with other strategies. AI as an adaption method was effective with six (6) of the eight (8) approaches proving successful. AI opponents have limited utility in serious games by themselves, as not all serious games may have opponents or competition-based game mechanics. However, Hunicke (2005) demonstrated that AI opponents, as part of a broader set of game mechanics, can be highly effective. AI opponents, similar to those described by (Demediuk, Tamassia, Li, & Raffe, 2019; Ilici, Wang, Missura, & Gärtner, 2012; Tan, et al., 2009), may be effective in human player versus computer simulations and serious games.

6.7 Direct Player Assistance

Player assistance was used in eight (8) articles (Baldwin, et al., 2014; Hunicke, 2005; Liu, et al., 2009; Negini, et al., 2014; Parnandi & Gutierrez-Osuna, 2014; Pato & Delgado-Mata, 2013; Stein, et al., 2018; Vicencio-Moreira, et al., 2014). In the experiments by Pato and Delgado-Mata (2013) and Vicencio-Moreira, et al. (2014), 2D and 3D shooting games respectively, players were provided direct assistance in the form of increased damage, aiming assistance and movement speed. However, in both cases they demonstrated mixed results. When those levels of assistance are expanded to also include environment assistance and opponent modifications, the systems were more successful at improving gameplay outcomes (Hunicke, 2005; Negini, et al., 2014; Stein, et al., 2018). This approach may thus be a valuable method for helping players in serious games and simulation training when combined with multiple DDA methods (Stein, et al., 2018).

6.8 Encounter Adjustment

This technique makes the challenge of the game fundamentally easier or harder, and aligns to the absolute difficulty of a task (Adams, 2014). This technique has similarities with environment content adaption, but differs in that it adjusts active aspects in the game, such as red lights and number of NPC cars (Prendinger, Puntumapon, & Madruga, 2014) rather than the static environment, or the number and abilities of enemies (Catarino & Martinho, 2019; Zook & Riedl, 2014). This technique was used in sixteen (16) studies, seven (7) of which used it as one of multiple methods (Abdessalem & Frasson, 2017; Alves, et al., 2018; Hunicke, 2005; Negini, et al., 2014; Papadimitriou, et al., 2019; Plass, et al., 2019; Tadayon, et al., 2018) and nine (9) as a single adaption method (Catarino & Martinho, 2019; Fernandes & Levieux, 2019; Ninaus, et al., 2019; Prendinger, et al., 2014; Suaza, et al., 2019; Van Lankveld, et al., 2009; Yun, et al., 2009; Yun, Trevino, Holtkamp, & Deng, 2010; Zook & Riedl, 2014). Single versus multiple adaption strategies were similarly effective with 78% success and 75% success respectively.

6.9 Environment Content Adaption

Environment Content Adaption refers to adding or removing content from game levels that can aid or hinder the player; this can include elements like health packs or ammunition, turrets and traps. Environment Content Adaption was used in five (5) papers (Alves, et al., 2018; Bian, et al., 2019; Hunicke, 2005; Negini, et al., 2014; Parnandi & Gutierrez-Osuna, 2014) already discussed, as one of several adaption strategies.

7 Summary of Suitability of DDA Approaches to Serious Games

This systematic literature review did not identify a clear singularly superior DDA method for games, or indeed, serious games and simulation training. The primary reason is that the best strategies are likely to be dependent on the type of game being designed. Nine (9) papers articulate that multiple adaption strategies may be beneficial over single adaption strategies (Avi Shena, Sitohang, & Rukmono, 2019; Bontchev & Georgieva, 2018; Burns & Tulip, 2017; El-Habr, et al., 2019; Liu, et al., 2009; A Nagle, et al., 2014; Parnandi & Gutierrez-Osuna, 2014; Yun, et al., 2009; Yun, et al., 2010). Additionally, Liu, et al. (2009) state that "…the next generation of DDA mechanisms should consider both player's performance and affective state information" (p. 508). This is reinforced by Parnandi and Gutierrez-Osuna (2014), who conclude that by tracking multiple metrics (performance, arousal and skill) a "control law" could be developed to "…dynamically maintain a proper balance between difficulty and skill level such that a player stays in flow" (p. 41).

Out of the 59 studies reviewed in this research, 21 measured some form physiological system of which 18 do this to infer an emotional state as the trigger for their DDA mechanism. These affect-based systems appear as effective as other non-physiological methods with a 68% success rate, however, they suffer some significant limitations outside of a clinical environment as they require additional equipment or devices to measure physiological signals. Additionally, physiological measures are prone to high variability inherent in the measure of human systems (Burns & Tulip, 2017). While this potential shortfall was noted in several studies, technology improvements will likely see further integration of affective technologies into training systems in the future.

To categorise how effective the different DDA systems were, we reviewed the data provided in each article and considered reported measures of success. For serious games and simulation training, success was considered as improvement in task performance, retention of learning material, or engagement whereas success for entertainment games is considered as greater enjoyment/engagement or better performance (e.g. scores) within the game. Any mixed, inconclusive, or caveated results are unsuccessful for the purpose of this analysis; only clearly effective DDA systems were given a positive rating. The results of this analysis (Table 4) were then subjected to four goodness of fit chi-squared tests. These tests were conducted in order to ascertain if there were any significant differences between single and multiple adaption strategies in 2D and 3D games. However, this does not consider the difference between the goals of serious games compared to entertainment games. From the data in Table 4, a goodness of fit chi-squared tests can be calculated using a p value of .05 for four null hypothesis:

H_0. A single 2D adaption (π_A) is equally effective as a single 3D adaption (π_B):

$$H_0 : \pi_A = \pi_B$$

$H_{1.}$ A multi 2D adaption (π_C) is equally effective as a multi 3D adaption (π_D):

$$H_1 : \pi_C = \pi_D$$

$H_{2.}$ A multi 3D adaption (π_D) is equally effective as a single 3D adaption (π_B):

$$H_2 : \pi_D = \pi_B$$

H_3. A multi 2D adaption (π_C) is equally effective as a single 2D adaption (π_A):

$$H_3 : \pi_C = \pi_A$$

Table 4. Comparison of single vs multiple adaption methods

	DDA	2D Game	2D success	3D Game	3D success	No. total Successful DDA	Percent total successful
Single Adaption	43	28	19 (68%)	15	11 (73%)	29	67%
Multiple Adaption	16	7	3 (43%)	9	8 (89%)	11	69%

H_0 (*p = .064*), H_1 (*p = .095*) and H_3 (*p = .154*) did not reject the null hypothesis. However, H_2 demonstrated a significant difference (*p = .037)*:

$$X^2 = (1, N = 24) = 3.84, p = .037$$

Therefore, from the papers reviewed for this study, in 3D games, multiple adaption strategies are likely to be more successful. This is relevant to serious games where multiple adaption strategies can be combined with the concept of multiple measures assessing multiple levels of competence, via a range of performance metrics and cognitive load. This idea is supported by Salas, Rosen, Held, and Weissmuller (2009), who indicate that "[P] performance diagnosis requires the integration of multiple measures from multiple data sources and at multiple levels" (p. 368).

8 Conclusion

In summary, we present the results of a systematic review of DDA approaches in serious games, simulation-based training, and entertainment gaming. The data collection conducted between February 2019 and September 2019 and again on 20 January 2021 used the formal PRISMA process (Moher, et al., 2015) to reduce 472 articles retrieved to a relevant set of 59 for this review. The majority (73%) of the approaches identified in the review applied single adaption approaches to 2D games (65%) and 3D games (35%). This single adaption approach to 2D games was 68% successful, and 73% successful when applied to 3D games. The remaining 27% of the articles reviewed applied a range of multiple adaption strategies to 2D games (44%) and 3D games (56%). Of these multiple adaption approaches, 3D games were 89% successful compared with only 43% successful in 2D games. A goodness of fitness chi-squared test was conducted identifying that there is a statistical difference in multi-adaption strategies in 3D games (p = .037). These findings indicate that a single DDA approach for 2D games may be suitable, however in more complex 3D games, a multiple DDA approach is more likely to succeed.

A multiple measure DDA approach may prove beneficial for serious games and simulation training as there is, on the face of it, more than one aspect that should be measured: the participant's performance relating to the pedagogical content, their performance or response to the game content and their level of engagement or cognitive load, which have been shown to correlate (Sharek & Wiebe, 2014). From a learning perspective, of the papers reviewed, only one utilised a measure of working memory (Afergan, et al., 2014), which is related to cognitive load. Cognitive load in serious games and training has been shown to be an indicator of success (Aldekhyl, et al., 2018; Woo, 2014), and therefore should be further explored within DDA systems applied to serious games and simulation training (Dideriksen, et al., 2018). Existing literature on learning theories points to their relevance in ascertaining performance and mastery, particularly Cognitive Load Theory, which is directly applicable to developing DDA for serious games and simulation training.

An effective serious game should improve educational outcomes and increase engagement. This increases the complexity of a DDA system for serious games, as the difficulty should adapt to both aspects. It is important when designing an adaptive serious game, to consider how it will fit into the curriculum, the pedagogical soundness of the adaption, and how the serious game will be deployed and played. Landsberg, et al. (2010) identified the need for quality adaption in serious games to maximise the potential benefits associated with cost-effectiveness and approximation to one-on-one tutoring. Serious games rarely have the budget or resources of entertainment games, and therefore should ensure that efforts to incorporate adaption mechanisms focus on those most likely to be effective (Peirce, et al., 2008).

Additionally, serious games and simulation training DDA systems should adapt learning content and game elements, drawing on recognized learning frameworks and concepts (Hamari, et al., 2016). Of the 23 serious games included in this review, eighteen (18) reported successful outcomes from the employed DDA approach (a 78.3% success rating), and fourteen (14) of these successful 18 referenced learning or psychological theories (77.8%). An approach using cognitive-affective theory of learning with media outlines how learning principles can be successfully applied to interactive multimodal learning content (Moreno & Mayer, 2007), and therefore we conclude that the likelihood of a successful application of comparable learning theories to DDA development, in serious games and simulation training, is likely to be similarly effective.

Cognitive load, combined with performance measures, can also clearly indicate mastery or competence and therefore may accurately trigger the dynamic adaption of learning environments. Further research considering multi-measures for DDA triggers, combined with multi-adaption DDA mechanisms, is warranted to better understand their effectiveness over single, trigger-based approaches. Ideally, in serious games and simulation training, these measures will indicate the mental effort being extended by the learner and contrast that with performance to assess whether the correct level of challenge has been achieved. Thus, additional research in adjusting both the pedagogical content and game-play content to provide a nuanced and flexible DDA system should be combined with a dual measure approach to ascertain if this provides an ideal matrix.

Despite adopting the robust PRISMA review method, the results are not without potential limitations, particularly in relation to the scope of the search terms, and the application of exclusion and inclusion criteria developed to isolate relevant articles. From a scoping perspective, no date restrictions were applied, however, the terms used may not have captured all relevant articles. Additionally, the results were obtained from primary academic bibliographic databases; articles not referenced within these databases will not be included, nor will information on any proprietary commercial DDA approaches. The search terms retrieved a large set of results that were reduced through semi-automatic and double-blind peer review based on inclusion and exclusion criteria. While these criteria are provided and allow for replication of the results, obviously, the use of different criteria would produce a differing set of relevant articles. However, the inclusion and exclusion criteria applied for this research produced a result set that directly addressed the aims of this research.

Finally, DDA is an important and growing consideration in serious games and simulation training. The systematic review results reported in this research lay a foundation for those seeking to design and develop effective, adaptive games for learning and training contexts. Also highlighted are the various DDA mechanisms that can be used to achieve these outcomes. Future research should consider the use of multiple adaption strategies, and the area of cognitive adaption is identified as a fertile area for further consideration and research.

9 Acknowledgment

This work was supported by the Defence Innovation Network NSW PhD Grant, and the University of Newcastle Postgraduate Research Scholarship.

Appendix A – Reviewed Studies

Paper	Game Type	Genre	DDA Success	Adaption Measure	Adaption Technique	Learning/Psych Theory Referenced	No. Participants	Brief Description
Abdessalem and Frasson (2017)	**3D** VR driving game, *AmbuRun*.	Serious Game	Yes	**Physiological**: Eye tracking and EEG to infer affect: frustration and excitement.	**1. Speed**: Vehicle velocity. **2. Encounter Adjustment**: Amount of traffic.	NIL	20	This study uses a 3D VR game in which the player drives an ambulance, the game adapts based on player emotion. In their method they preserve a person in a desired emotional state.
Afergan, et al. (2014)	**2D** UAV flight path serious game.	Serious Game	Yes	**Physiological/Cognitive Load:** Functional near-infrared spectroscopy (fNIRS), measured working memory.	**1. Task Complexity**: UAV's added or removed.	Flow	12	This study demonstrates how fNIRS can be used to measure working memory and adapt task complexity in a simulated UAV flight path management serious game, this led to a 35% error reduction, compared to a control group.
Alexander, et al. (2013)	**2D** Side scrolling 'Flappy Bird' style game – *Star Surf*.	Entertainment	Yes	**Performance**: distance travelled while collecting objects.	**1. Procedural Content Generation**: obstacles, objects and gates. **2. Speed**: rate at which the game scrolls.	Flow	90	This article uses a 2D side scrolling game and compares the efficacy of static vs DDA difficulty setting for casual and experienced player types. They found that players consistently chose difficulty settings that were not suited to their playing style which may lead to boredom or frustration by having made the game too hard or easy.
Alves, et al. (2018)	**3D** First Person Shooter (FPS).	Serious Game (Health)	No	**Physiological**: Affect via EEG and heart rate.	**1. Encounter Adjustment:** enemy speed, health and spawn rates, player speed and shooting skill. **2. Environment Content Adaption:** number of health and ammunition pick-ups.	Flow	21	This study aims to achieve a flow state. They developed a performance based version as a control and they incorporated three tiers of difficulty. Every 20 seconds the DDA system adapted difficulty. The performance based system outperformed the affective system.
Andrew, et al. (2020)	**3D** First person horror.	Entertainment	Yes	**Physiological:** Measures valence and attention via facial expression recognition.	**1. Speed:** adjust the figure movement index (FMI) based on attention and emotion.	Flow	31	In this article the experiment seeks to compare emotion based DDA against a performance based system. They measured emotion via facial expression recognition, and the success was measured

								with a satisfaction survey combined with performance metrics. 61.29% preferred the emotion based DDA approach.
Avi Shena, et al. (2019)	**2D** Maths quiz game where player fights a dragon.	Serious Game	Yes	**Performance:** Time taken to answer plus current player score (correct answers) vs enemy power and speed. They have minimum and maximum values based on prior exam results and this matrix is used to adjust the DDA.	**1. Time:** Player is given more or less time to solve maths puzzles.	Flow	39	This article explores a DDA approach based on evidence-centered Design (ECD). The system was successful, claiming a 67% win rate which achieved the target goal. They use their ECD-based system to measure what is essentially competence. They comment that an additional adaption/guidance (pedagogical agent) would be ideal to support the learning content.
Baldwin, et al. (2014)	**3D** FPS.	Entertainment	No	**Performance**: Player score, difference between player and leader score.	**1. Direct Player Assistance:** Provide weaker players with a bonus to 'shield' points.	SDT Flow	62	This study explores how players' respond to knowing if a DDA system is being used. They used GSR as an objective measure of arousal and the Player Experience of Need Satisfaction (PENS) survey to measure player affect. The DDA mechanism improved the performance of weaker players. When players were aware of the DDA the efficacy was reduced, this was attributed to better players adopting strategies to offset the DDA.
Bian, et al. (2019)	**3D** VR Driving game.	Serious Game	Yes	**Physiological:** Biopac MP150 to measure GSR, respiration and heart rate in order to ascertain engagement. **Performance:** Success in driving task (obeying rules of the road), difference between current speed and speed limit, steering wheel and accelerator pedal manipulation.	**1. Task Complexity:** Changes to brake, accelerator and steering responsiveness. **2. Environment Content Adaption**: intensity of light in the environment. **3. Speed:** the speed of other drivers on the road.	Flow	20	This article explores how VR can be used to assist individuals with ASD learn to drive. The research combines extensive performance metrics with physiological; sensors as a combined measure with multiple adaption strategies. The comparison was between a performance only system versus a combined system (engagement and performance). The latter version was an improvement in all regards, and proved to be more effective than performance only measures.

Bontchev and Georgieva (2018)	**3D** First person action puzzle adventure game *Rush for Gold.*	Serious Game	Yes	**Physiological:** Facial tracking & galvanic skin response for emotion. **Performance**: Shooting, discovering and solving puzzles.	**1. Task Complexity:** Hidden items are made more difficult to find and puzzles made more complex. **2. Speed:** Flying gold bricks move faster or slower.	Kolb's theory of experiential learning	30	This article explored playing style recognition in an affective adaptive game. Facial tracking was moderately effective method used to infer affect however the method suffered from a number of shortfalls. They indicate that player style, and affect, should be combined with other measures in order to create effective DDA. Their player style recognition algorithm returned an 82.71% accuracy rating.
Burns and Tulip (2017)	**2D** Side scrolling 'flappy bird' style – *Star Surf.*	Entertainment	No	**Physiological**: COTS software combined with a web cam for facial recognition measuring emotion.	1. **Task Complexity:** Adjusts gap width in each 'gate'. **2. Procedural Content Generation:** interval between gates adjusts. **3. Speed**: gate speed across the screen.	Flow	15	This study found that players often adopted a flat facial aspect, effectively hiding their emotions, which reduced the effectiveness of the system. They suggest a "hybrid system" incorporating performance metrics may be more effective.
Catarino and Martinho (2019)	**2D** Constant wave battle game.	Entertainment	Yes	**Performance:** measure mastery based on challenges, obstacles and 'tags'. Including number of taps by player compared to the number designated for that difficulty. Also tracks content in order to vary it for interest.	**1. Encounter Adjustment:** Generate waves of enemies in differing formations, with different enemies.	Flow	32	They developed a player progression model that measures player performance and game repetition in order to both vary difficulty and content. The system compares player performance with the last 10 obstacles to predict the next wave. Their system was successful in increasing game play-time (number of games as well as game duration) which was their aim.
Chanel and Lopes (2020)	**2D** *Tetris.*	Entertainment	No	**Physiological:** measures electrodermal activity (EDA) to measure boredom and anxiety levels.	**1. Speed:** Speed of the Tetris pieces.	Flow	56	They measure EDA in order to ascertain player affect. A model was trained using deep neural networks prior to the main experiment. The goal was to measure boredom and anxiety and oscillate between those thus attaining Flow. This measure was a classifier result measured for periods of 20 seconds and decisions about adjustments made every 10 seconds. They compared two approaches to affect-based DDA, an absolute and a relative version. Neither worked particularly well.

De la Cruz, et al. (2018)	**2D** Quiz adventure serious game *Bacteria Madness.*	Serious Game	Yes	**Performance**: Quiz answers.	1. **Task Complexity:** Different questions were presented depending on player answers.	NIL	36	This article details a simple 2D quiz adventure game to help teach microbiology. They mapped out the curriculum and developed a probabilistic model based on a Bayesian network of the course, and a structure was generated mapping this out. This model predicted player performance and presented different questions based on this model, helping guide the students to success. The DDA version of Bacteria Madness was more successful than the non-adaptive version.
Demediuk, et al. (2019)	**2D** Fighting game called *FightingICE*.	Entertainment	Yes	**Performance**: Player health vs enemy health combined with available actions.	1. **AI:** Utilised a system called a Reactive Outcome Sensitive Action Selection (ROSAS) which uses MCTS-UCT.	Flow	31	This study set out to test a variety of AI based DDA systems, it utilised a Monte-Carlo Tree Search algorithm. The method that was shown to be most effective was an AI agent that had a worse win percentage than 50%. This highlighted that not all players seek an even match and may play games for a range of reasons, and that an even challenge is not always ideal.
Denisova and Cairns (2015)	**3D** 3^{rd} person Shooting game, based on Unity game *Nightmares*.	Entertainment	Yes	**Performance**: Damage inflicted.	1. **Time:** the time allotted to achieve a score was subtly adjusted to allow more or less time based on performance.	Flow	42	This article produced the only paper in which time is the principle adaption mechanism. They subtly manipulated the in-game time counter to count-down quicker or slower depending on player performance in order to increase difficulty and pressure. Their results indicated that the players were unaware of the adaption and this adaption led to a better player experience in a 3D 3^{rd} person shooter cartoon style game.
El-Habr, et al. (2019)	**2D** *Runner* a 2D endless platform runner game.	Serious Game (Health)		**Performance:** Measured player deaths as a result of trap platforms.	1. **Procedural Content Generation:** to generate trap frequency and yellow coin platform frequency.	Flow	10	The system used trap platforms and yellow coins to encourage the player to jump and crouch towards better health and fitness. The efficacy was measured by score, calories burned, balanced use of body parts (e.g. left and right arm), and variability of body movement. The system was successful in increasing game score and calories burnt. They discussed possibilities to further adapt based on measuring player movements and calories etc. This would represent a dual

								measure approach combining performance with physiological activity.
Ewing, et al. (2016)	**2D** *Tetris.*	Entertainment	No	**Physiological**: They used an EEG to measure mental workload and boredom.	1. **Speed**: the speed of the falling pieces was altered.	Flow	10	In this paper the NASA-TLX questionnaire is used to measure subjective workload. The final experiment tested three variations of DDA with one control type. The control allowed the player to set the Tetris fall speeds, and this player choice was taken as a preference measure. The DDA system either over- or under-adapted the fall speeds compared to player preference, indicating the system was not successful.
Fernandes and Levieux (2019)	**2D** Top Down tank arcade shooter.	Entertainment	Yes	**Performance:** They measure “success” defined as player win or die ratio.	1. **Encounter Adjustment:** they adjust number of enemies, enemy movement speed, enemy turn speed, enemy time.	Flow	37	The experiment initially uses a simple adaption method until there are enough data points at which time they switch to a logistic regression algorithm as the difficulty value. They then use this value to adjust number of enemies, enemy movement speed, enemy turn speed, enemy time between shot, and enemy accuracy. The algorithm was successful however seemed to suffer a very rapid increase in difficulty (more exponential than linear) from very easy at 0.4 to very hard at 0.6. There was no assessment of player enjoyment reported or engagement.
Hendrix, et al. (2019)	**2D** Side scrolling platformer.	Serious Game	No	**Performance:** measured distance travelled compared to time expired.	1. **Procedural Content Generation:** to generate and place the platforms for the player to jump to. Specifically adjusting platform size and platform gap.	Flow	37	In this article the authors tested two video games, one an existing game with a convenience sample of eight (8) players and one with 37 participants built from scratch by the authors. Due to the small convenience sample and pre-made nature of the first game, this review regards the second custom made game with the larger sample when assessing success. This is based on the assumption that greater effort and rigour went in to that

								experiment (as evidenced by the additional effort typical for making a game from scratch) and the larger sample size indicating priority. Analysis of their results indicated no statistical significance towards "fun" for the adaptive version.
Hooshyar, et al. (2021)	**2D** Computational Thinking puzzle game.	Serious Game	Yes	**Performance:** They use a Bayesian Network to measuring commands used to solve problems combined with situation in the game, score, step life and more. And they classify solutions presented.	1. **Task Complexity**: the cat is moved to more or less challenging locations for the puzzle. 2. **Feedback:** hints, feedback and links to learning tutorials are provided.	NIL	79	In this paper the authors used a game called Autothinking that improve computational thinking skills and conceptual knowledge. The game used a Bayesian network to "connect observable game activity with a learning outcome". Measuring commands used to solve problems combined with situation in the game (mouse, cats, cheese), score, step life and more. And they measure solutions presented (i.e. whether they are satisfactory, good, or unsuccessful) and provide appropriate hints, feedback and links to tutorials. They also provide adaptions to gameplay by moving a game cat to different zones. This game also improved general attitudes to learning computational thinking skills in addition to the skills themselves.
Hunicke (2005)	**3D** FPS, *Case Closed.*	Entertainment	Yes	**Performance:** Estimates the probability of damage likely to be suffered by the player and adapts accordingly. If the system determines chance of death is greater than 40% it starts to intervene.	1. **AI:** The 'Hamlet' system itself and enemy AI settings (class). 2. **Environment Content Adjustment:** health, ammunition, weapons etc. availability. 3. **Player Assistance:** modifying player hit points, strength of their attacks and accuracy. 4. **Encounter Adjustment:** hit points, weapons and location.	NIL	20	The AI system, Hamlet, determines the probability of death in any given encounter and then intervenes based on one of three intervention strategies, these strategies influence inventory items for the player (health, ammunition and weapons) and also player hit points, accuracy, strength of their attack and more. Hunicke's system reflected the concept of aiding the player to achieve without negatively impacting the game experience through a range of subtle interventions.
Ilici, et al. (2012)	**2D** Board game, Chinese Chess.	Entertainment	No	**Performance:** Measure player moves over time and develops a prediction for correct difficulty level.	5. **AI**: opponent in checkers and Chinese chess. This AI uses a "partially ordered set master (POSM) algorithm" (page 57).	Nil	12	The algorithm in this article establishes a belief for appropriate difficulty based on player moves over time, as the game progresses the correct difficulty level is achieved. This is evaluated based on a

								minimax algorithm termed POSM. POSM encountered issues with players who played lots of random moves, during which the human players were unable to win a single match, however in most other scenarios POSM was successful.
Li, et al. (2014)	**2D** Side scrolling.	Entertainment	No	**Performance**: Difficulty is adjusted after 15 objects are passed.	**1. Procedural Content Generation:** they used a pattern generation system to procedurally place game obstacles. **2. Speed:** the speed of the obstacles adjusts as they fall down the screen.	Flow	30	This study tested a DDA version of a game against a static difficulty system and found no statistically significant performance difference between the two methods across 30 participants. However they suggested that this may be due to short game times.
Liu, et al. (2009)	**2D**: *Pong*.	Entertainment	Yes	**Physiological**: "[…] various cardiovascular, electrodermal, electromyographic and body temperature signals" (page 509). They measured anxiety as the emotional state.	**1. Task complexity:** ball size, keyboard responsiveness, **2. Player assistance:** Paddle size **3. Speed**: ball speed and paddle speed	Nil	9	This study compares a performance based adaption with affective adaption of the game Pong. Their system used a regression tree based algorithm for affective modelling, their system, in real-time, attained a 78% accuracy when compared to player self-assessment. The difficulty system had three settings in both performance and affective versions. 6 of 9 participants recorded lower anxiety levels with the affective system and 7 showed greater improvement in performance with the affective system. Liu et al. (2009) acknowledge that the physiological sensors required may be "…restrictive under certain circumstances".
Lora Ariza, et al. (2019)	**2D:** *Tetris*.	Entertainment	No	**Performance:** measures current game performance and compares it to an offline profile of good games.	**1. Task Complexity:** Compares current score, board height and holes to previous cases seeking a match for "good games". If player is playing worse the game is made easier.	Flow	26	This game uses a dual measure with case based reasoning of player profile and current game state. Each case represents 10 Tetris pieces and measures: number of the piece since start of game (time), score, number of holes (empty spaces), board height,. An offline database of good games is stored: At the end of a game the player is asked to rate the game 1-5, a score of 4-5 = a "good game". They compare this concept of a good game to a series of offline cases and adjust the game accordingly by modifying the choice of Tetris pieces. They saw no statistically significant improvement.

Mainieri, et al. (2018)	**2D:** 2.5D maths game with a penguin traversing an environment.	Serious Game	No	**Performance:** the game measures if the player successfully answers a question, skips it or answers it using the rewind button.	**1. Task Complexity:** results are used to adjust a variable called the "Success Streak". This then presents the next question/challenge as either of the same difficulty or increasingly challenging.	Flow	10	This paper explored creating an arithmetic game for primary school children. This initial experiment was only trialled on 10 educators (not the ultimate demographics) and had other limitations. While some understandings were reached there was no conclusive outcome, yet.
Masanobu, et al. (2017)	**2D:** *Tetris*.	Entertainment	Yes	**Performance:** The game adapts based on the Tetris stack height and the variation in height of the adjacent blocks.	**1. Speed**: the speed of the falling Tetris pieces increases or decreases. **2. Task Complexity**: The shape of the Tetriminos is categorized into best, bad and worst. Based on stack height the appropriate tetrimino is selected for the desired outcome.	Flow	23	The article looked at extremes of adaption, from very easy to intensely difficult, and they termed this concept Dynamic Pressure Cycle Control (DPCC). The DPCC system cycles between difficulty to adjust the sense of pressure and fun, this is equated to the fun of a roller coaster ride. This DPCC version was tested against a control were the difficulty adjusted purely randomly, the DPCC version was strongly preferred. This experiment demonstrated that short bursts of exceeding difficulty may increase enjoyment in a game when counter balanced by period of relative calm.
Moon and Seo (2020)	**2D:** Airhockey.	Entertainment	Yes	**Performance:** Measures player win loss ratio and movements and this is compared to an offline model previously generated.	**1. AI:** Adjusts the AI system to try and be similar to the player, and also adjusts the difficulty up or down depending on player win or loss ratio.	NIL	9	The authors explored using a "fast user adaption method" requiring only a small amount of player data, utilising a meta-learning algorithm. Their system outperformed other deep-learning methodologies and had a much faster learning speed.
A Nagle, et al. (2014)	**3D:** First Person Memory serious game.	Serious Game (Health)	Yes	**Performance**: was based around the accuracy of placing the objects in the correct location as well as the number of times the list was viewed.	1. **Task Complexity:** adjusted by the number of objects, and the number of times the player could view the list of those objects and locations they are to be placed in.	ZPD Flow SDT	24	In this study a game was designed to test the efficacy of automatic performance based adaption against player controlled adaption. Specifically, they assessed the efficacy of SDT and Flow/ZPD concepts to ascertain if either was more effective in a serious game. Both versions were effective, but they closed with an interesting opinion: "…that a "trade-off" between maximizing performance and maximizing enjoyment is needed to ensure both short-term and mong-term learning in serious games." (Page 127).

A. Nagle, et al. (2015)	**2D**: Memory training serious game.	Serious Game (Health)	Yes	**Performance**: was measured based on the number of correct and incorrect memory sequences remembered.	1. **Task Complexity:** three good rounds means performance rated as "Good", two bad rounds were rated as "bad". If "good" the memory sequence increases by 1, if "bad" it decreases by 1. Minimum difficulty was 3 and maximum 7.	Flow	14	This article assessed two versions of the game, both had DDA but one included visual art changes in the game, background, foreground and animation. They showed that changing the art elements led to greater enjoyment and attention when compared to the non-visual enhanced DDA version.
Negini, et al. (2014)	**3D**: FPS game.	Entertainment	Yes	**Physiological:** measured galvanic skin response (GSR) to infer player arousal.	1. **Player Assistance:** speed and access to weapons 2. **Environment Content Adaption:** rate of health packs spawned and visibility via fog 3. **Encounter Adjustment:** their speed and quantity	SDT Flow	16	The article explored what type of adaption was best suited for affective games. They explored three methods of DDA. Their system was successful and showed that supporting the player to achieve mastery of the game was more effective than making the enemies weaker. This correlates with the concept of mastery in SDT (Richard M Ryan & Deci, 2017; Richard M. Ryan, et al., 2006).
Ninaus, et al. (2019)	**3D:** Emergency management game.	Serious Game	Yes	**Physiological:** measured heart rate, if HR dropped below base rate for 10Sec then difficulty increases, if heart rate increases more than 5bpm over base, then the game is made easier.	1. **Encounter Adjustment:** the number of tasks a player has to deal with simultaneously is adjusted up or down. E.g. extra bystanders would spawn and get injured to make the game harder. To make it easier a helicopter may appear to help transfer injured people.	Flow	12	This serious game revolved around the management of an emergency casualty event. The authors compared adaptive vs non-adaptive versions of the game. The experiment measured heart rate and adjusted based on perceived levels of stress, difficulty was principally adjusted by dealing with the volume of cases a player needed to deal with. The results were positive indicating success.
Pagalyte, Mancini, and Climent (2020)	**2D:** Turn based battle game.	Entertainment	No	**Performance:** the system measures the number of losses and wins in a row and adjusts enemy performance to match the desired result based on 17 different agent "states". The main focus is 5 wins or losses in a row.	1. **AI:** They use a reinforcement learning approach to modify the agents (NPCs) choice of actions.	Flow	10	The authors explored a reinforcement learning approach with a State-Action-Reward-State-Action process. This process adjusted the choices the enemy makes from five different moves, incorporating an Exploration vs Exploitation algorithm. Exploration the agent does random moves and in exploitation the agent chooses the most effective move. The aim was to see if the reinforcement learning agent-based DDA engendered a "better" gameplay experience. There was no statistically successful outcome.

Papadimitriou, et al. (2019)	**2D:** Educational adventure serious game to learn HTML programming.	Serious Game	Yes	**Performance:** fuzzy logic system combined with quiz results measure player knowledge level. Four "fuzzy sets" are defined: Unknown, Unsatisfactorily Known, Known and Learned.	1. **Feedback**: Pedagogical agents provide feedback and assistance where necessary, they also apply quizzes. 2. **Encounter Adjustment**: increased the number of level sections and NPC providing quizzes.	NIL	60	In this study a 2D educational adventure game was developed with pedagogical agents that provide direction and feedback, including quizzes that can lead to additional game content. The main adaption mechanism is activated after the player fails the various quizzes then additional rooms and puzzles are added to the game in order to reinforce the player understanding. The authors incorporate an assessment of mastery through a student modelling tool which includes a fuzzy inference system. The game was successful in four areas: usability, likeability, educational effectiveness and adaptivity.
Park, et al. (2014)	**2D**: Music rhythm game.	Entertainment	No	**Physiological**: they used an EEG to measure emotional states: boredom, anxiety and engagement.	1. **Task Complexity**: the number of notes is increased from 74, to 159 and finally 280 notes, the speed is unchanged.	Flow	8	This article used a system that required the participants to play the game at a fixed incremental difficulty setting first in order to establish a baseline for the emotion detection. The paper states that "[…] the performance of this system is not competent", suggesting that Park et al. were unsuccessful in developing an effective DDA system.
Parnandi and Gutierrez-Osuna (2014)	**3D**: Driving game in which player only controls steering.	Entertainment	Yes	**Physiological**: electrodermal activity (EDA) use to measure arousal.	1. **Task Complexity:** the velocity of the player car is adjusted 2. **Player Assistance:** steering is adjusted by the addition of a random noise disturbance. 3. **Environment Content Adaption:** weather is used to adjust visibility distance	Flow	25	This paper explored a 3D driving game that utilises electrodermal activity to inform the DDA mechanism that then adjusted the steering, speed elements and weather/visibility. In the experiment car velocity was found to be the most effective adaption technique.
Pato and Delgado-Mata (2013)	**2D**: one-vs-one player shooting game.	Entertainment	No	**Performance**: the number of hits is compared as well as the time between successful hits.	1. **Player Assistance**: player movement speed and damage per shot is adjusted.	Flow	26	This article focussed on a DDA system to aid players in a one-vs-one 2D shooting game. The algorithm employed measures the number of hits and the time between hits, then adjusts the weaker players speed and damage done per hit. They reported inconclusive results expressing a need for further experimentation.

Peirce, et al. (2008)	**3D**: First person adventure serious game, called *ELEKTRA*.	Serious Game	No	**Performance**: based on task completion about the physics of optics for high school students.	**1. Feedback**: pedagogical agents provide feedback, hints and tips to help players succeed.	Flow	49	The study was the only one where the distinction between educational content and game experience is articulated and that the two can operate independent of each other. They sought to develop a system of adaption that may be applicable, and reusable, across multiple games. The adaption strategies within ELEKTRA are "[…] (1) motivational and hinting support, and (2) meta-cognitive feedback." The main feedback mechanism was provide through an NPC called Galileo. Their experiment showed promising results, though not statistically significant and has prompted further study.
Pfau, et al. (2020)	**3D:** MMORPG Game *Aion.*	Entertainment	Yes	**Performance:** They used a Deep Player Behaviour Modelling (DPBM) system to model game state based on information about the player, opponents and skill history in order to determine actions (skill usage) probabilities.	**1. AI:** the DPBM adjusted skill usage in order to attempt to emulate the player.	Flow	30	A live test was run on a server in the MMORPG game Aion, the authors demonstrated their dungeon, with two types of DDA (one being the DPBM), was more successful than any non-DDA instance. Their DPBM DDA application was more successful than the alternative tested approach. This was measured based on player motivation. Their system attempted to match the player skill to provide the ideal level of challenge.
Plass, et al. (2019)	2D: A variation of a previously validated serious game called *Alien Game.*	Serious game	No	**Performance**: after three correct feeding actions the game difficulty goes up, if the player fails to feed an alien the correct food item or an alien isn't fed before disappearing off the screen the difficulty goes down.	**1. Speed:** speed the aliens move at. **2. Encounter Adjustment:** number of aliens and their associated rules.	Nil	75	In this study a 2D serious game was designed to help children train their executive function skill of 'shifting'. This experiment tested the non-adaptive version of Alien Game with a version with DDA. The results of the experiment did not reveal a significant difference between the adaptive and non-adaptive control, both were equally effective in improving executive function outcomes.

Prendinger, et al. (2014)	**3D**: Car driving game.	Entertainment	No	**Performance:** based on how economically the player is driving.	1. **Encounter Adjustment:** Frequency of red lights and number of NPC cars; through a distributed constraint optimization (DCOP) method	Nil	30	This study found that the assessed DDA system was unable to create a suitably difficult challenge, one reason proposed is that the actions available to the DDA system were too limited to create a hard enough challenge.
Reidy, et al. (2020)	**3D** VR cognitive training set in realistic environments.	Serious Game (Health)	Yes	**Physiological:** facial electromyography (EMG). Every 45 seconds the EMG classifies player affect. **Performance**: Based on correctly finding a list of items from a list.	1. **Task Complexity:** The number of items required to be found increased or decreased.	Flow	6	The authors explore cognitive training in elderly adults in order to help improve memory and cognitive function. The game was set in two VR environments and measured affect as well as performance to create a dual measure that informed a simple task adaption strategy. The game was successful in creating competence and possibly increasing intrinsic motivation.
Rietveld, et al. (2014)	**2D** Top-down racing car game.	Entertainment	Yes	**Performance:** multi-class Support Vector Machine (SVM) that calculates a measure based on the percentage of time the player is off-track, and the normalised average speed compared to an attainable maximum.	1. **Procedural Content Generation:** Track sections of varied types generated	Nil	11	In this article a driving game adapted the track by adjusting the sharpness of curves and bends, making them smoother and easier for underperforming players, and sharper harder curves, turns and bends for well-performing players. They compared this method against a technique termed rubber-banding (Tomlinson & Melder, 2013) and found their implementation was successful in achieving the correct level of difficulty.
Sampayo-Vargas, et al. (2013)	**2D**: Word/language serious game – Spanish cognate bubble game.	Serious game	Yes	**Performance**: Based on successfully answering language questions correctly, three correct made it harder and vice versa.	1. **Task Complexity:** 11 levels of difficulty defined by the number of boxes, no. of bubbles, bubble speed and the difficulty of the educational content.	ZPD Flow CLT	234	The game in this study included adaption of educational content but also the number and speed of game elements. Sampayo-Vargas, et al. (2013) express that DDA works best with multiple adaption methods, and expressly articulated that their game applies the concept of scaffolding. The DDA version of their game

								showed a statistically significant improvement in learning outcomes over two other tested methods.
Silva, Silva, and Chaimowicz (2017)	**2D**: Player-vs-player MOBA game, *Defence of the Ancients.*	Entertainment	Yes	**Performance**: An amalgam of Player Level, Player Death and towers destroyed.	**1. AI:** AI is developed with three levels. It determines the development pace of the player and adapts the AI performance to meet the player more closely.	Flow Malone Motivation Theory	11	This study created an AI opponent that established three varied levels of enemy difficulty that were then dynamically adjusted based on player performance. The AI player adjusted both defence and attack strategies of enemy players and was shown to be effective.
Spiel, et al. (2019)	**2D:** *Tetris*.	Entertainment	No	**Performance:** They use the overall pile height on the highest level as well as the difference of height between rows (also called bumpiness) to analyse a player's performance. **Physiological:** Eye movement in order to infer a level of competence.	**1. Speed:** adjust speed of falling pieces based on pile height and bumpiness.	Flow	43	In this experiment the authors compared their eye gaze and performance DDA method against a performance only DDA method. The eye gaze measure is added to the performance calculation to infer competence based on eye transition values which shows how players look at the game. There was no significant difference between the two systems.
Stein, et al. (2018)	**3D**: 3rd person shooter based on *Boot Camp.*	Entertainment	Yes	**Physiological**: EEG to measure affect	**1. Player Assistance**: speed increase, damage reduction, and weaker player invisible for a period and AI controlled turrets target stronger players.	Flow	24	This article compared the DDA system to a heuristic DDA system and a control with no DDA. Their results clearly indicated that their DDA method was preferred by players and achieved better results than the control and Heuristic DDA. However they made a comment that the EEG headset "[…] needs improvement if it is to be used in the gaming industry" (page 23).
Suaza, et al. (2019)	**2D:** Fighting game *The Forest of the Guardians.*	Entertainment	Yes	**Performance:** they measure player "aggression" based on the difference in HP of player vs opponent.	**1. Encounter Adjustment:** a heuristic function is used to pick a difficulty of easy, med, or hard in a finite state machine. The difficulty level changes op-	Flow	19	The authors explored a 2D fighting game called *The Forest of the Guardians* where player aggression is inferred based on the difference between player and enemy health. This difference drives the difficulty level with the system trying to have the NPC fighter with the same

					ponent: attack speed, movement speed and armour points.			health as the player. The game was considered balanced and thus successful.
Tadayon, et al. (2018)	**3D**: Physically active fruit slicing game for at-home motor learning.	Serious game (health)	Yes	**Physiological**: Facial tracking using a Microsoft Kinect camera towards the goal of achieving a flow-state, estimated using Visage emotion-vectors. **Performance**: The player is given an "Intelligent Stick" device and instructed to use it to slice the virtual fruit. How well they perform this action is recorded with a number of criteria (e.g. trajectory, speed and posture)	**1. Speed**: of fruit across screen **2. Encounter Adjustment:** size and number of fruit	Flow ZPD	1	This Article identified that there are problems with facial recognition as the system may be obscured by sunglasses or facial positioning. Their experiment has three performance based adaption methods (Bayes-Net, Cluster and Hit-Rate) each of which tracks player emotion state as an additional adjustment vector. The Bayes-Net adaption technique was the most successful of the conditions tried, Tadayon et al. (2018) found that combining their flow-state adaption with the three tested methods yielded "[…] affective game design that is more responsive to an individual's state in at-home unsupervised training."
Tahai, et al. (2019)	**3D:** Proportional reasoning game for children, *Scalebridge*.	Serious Game	Yes	**Performance:** 'win' or 'loss' dependant on successfully balancing the bridge within 60 seconds. **Physiological:** Neurosky MindWave EEG device to measure focus.	**1. Task Complexity:** there are 60 levels, gradually getting harder, the DDA enables the player to progress more rapidly and repeat levels as required. **2. Feedback:** a video guide is used to provide feedback.	Flow	20	This article describes the game *Scalebridge* in which a dual measure of focus and performance are combined to influence two adaption mechanisms. The states identified were: focussed/winning, focussed/losing, unfocussed/winning, and unfocussed/not winning which were used to inform the DDA mechanisms. The authors were seeking to test both flow and the speed of mastery, their DDA implementation was effective.
Tan, et al. (2009)	**2D**: Board Game – *Gomoku.*	Entertainment	Yes	**Performance**: the value of each move is evaluated, it is compared to the moves not taken. After every 9 moves, the level of difficulty is increased if the player made 7 or more 'best' moves during that turn. If the player made less than 3 'best' moves, the difficulty is decreased.	**1. AI**: heuristic AI with 8 difficulty levels, starting at level 4 and after each set of 9 moves the AI adjusts the difficulty based on the heuristic algorithm.	Nil	50	In this study an adaptive AI opponent, was made, incorporating a minimax algorithm and alpha-beta pruning. This type of AI system has limited use outside of strict rule based game-play, like Gomoku or Chess, and is not likely suited to more organic types of games with fast action and many diverse possible actions (Mayefsky, Anene, & Sirota, 2003). Minimax algorithms are used in two player games further limiting the use in serious games and simulation training (Mayefsky, et al., 2003).

Tang, et al. (2020)	**3D:** First person road network engineering game – *Gridlock*.	Serious Game	Yes	**Performance:** a knowledge-based classification is implemented within gridlock through a kNN-based closed-loop control. There are three modules: student, expert and pedagogy.	1. **Feedback:** The system makes the student look through learning material in the knowledge areas the system thinks the student lacks. Also it adapts through prompts and cues.	NIL	38	The game *Gridlock* collects information on the player performance via a k-nearest-neighbour (kNN) classification and provides feedback to the student in real-time. There are three modules in the game student, expert and pedagogical. The latter recognizes when the student is having difficulty and provides hints etc. The DDA implementation showed significant improvement and success.
Van Lankveld, et al. (2009)	**2D**: Role playing game – *Glove.*	Entertainment	Yes	**Performance**: player health vs. the average amount of damage enemies currently in the game can cause	**1. Encounter Adjustment:** the game spawns one of the three types of enemies at different rates.	incongruity theory	24	This article frames their system through the lens of incongruity theory referencing Rauterberg (1995). Van Lankveld et al. (2009) found some of their hypothesis relating to the relationship of DDA and gaming to Incongruity theory were satisfied but others were not. Their experiment results were mixed, however there were some positive correlations indicating DDA was moderately successful.
van Oostendorp, et al. (2014)	**3D**: 3D medical triage game - *Code Red Triage.*	Serious Game	Yes	**Performance**: Based on 4 criteria related to how well player triaged victims. The game has 6 tiers of triage challenge – if the player does well they move up a tier, if they don't do well they stay at the current tier until they master that level.	1. **Task Complexity:** the complexity of the medical case to triage was increased.	Flow	28	This study developed a 3D medical triage game. The game sought to ensure an appropriate cognitive balance to prevent either under-load or over-load of cognitive faculties, by enabling players to finish quicker or slower depending on their proficiency in the learning material.
Vicencio-Moreira, et al. (2014)	**3D**: First Person Shooter.	Entertainment	No	**Performance**: Subtracts the lower player kills from higher player kills to determine a level of assistance from 0-10.	**1. Player Assistance:** aiming assistance via bullet magnetism and area cursor.	Nil	20	The methods explored in this article showed no improvement in scores or enjoyment and the article concluded that DDA requires more than simply "helping weaker players aim".

Xiang, et al. (2013)	**2D**: *Tetris.*	Entertainment	Yes	**Physiological**: A facial recognition system based on the Active Shape Model (ASM), this system detects four expressions: frustration, excitement, relaxation and boredom.	**1. Speed**: of the falling Tetris blocks.	Flow	20	This study used an emotion-based DDA system that was tested against a performance based system and was preferred by 14 of 20 participants. The only adaption was the speed of the falling Tetris blocks.
Yun, et al. (2009)	**3D**: 3rd person robot shooting game – *Robot Game*.	Entertainment	Yes	**Physiological:** A thermal imaging system, called StressCam, to monitor stress in a players face via supraorbital skin temperature.	**1. Encounter Adjustment:** The game difficulty level adjusts once a minute, by making enemy robots stronger or weaker relative to the player.	Nil	14	The system in this article utilises two computers, one for the game and the other for managing the StressCam system. Prior to undertaking the game players completed a demographic and gamer survey in order to categorise the players in to one of four player types, this was followed by a three minute relaxation session to establish a StressCam baseline. The system increased player performance successfully, the system requires two computers to work.
Yun, et al. (2010)	**3D**: 3rd person robot shooting game – *Robot Game*.	Entertainment	Yes	**Performance:** Based on enemies killed, enemies encountered, player health and player deaths. This is then weighted by player type.	**1. Encounter Adjustment:** The game difficulty level adjusts once a minute, by making enemy robots stronger or weaker relative to the player.	Nil	57	This study explored player types and motivations as a foundation for a DDA system, they developed a Profile-based Adaptive Difficulty System (PADS). This system groups players by a profile type by asking two questions, these answers are used to tune the DDA mechanism. Similar to Yun et al. (2009) they used the same 3rd person robot shooting game with the same adaptions however in this paper DDA is based on performance rather than affective state. This system was successful and demonstrated the importance of considering player motivation as a lens through which to adapt games.
Zook and Riedl (2014)	**2D**: Turn-based spell battle RPG game.	Entertainment	No	**Performance**: The DDA measures player performance through a tensor factorization approach, modelling the player and predicting the player responses to future actions and then selecting game content to match the performance curve.	**1. Encounter Adjustment:** The DDA mechanism selects four enemies at a time that combined have a "predicted average performance value". Thus different enemy combinations create different difficulty levels.	Nil	30	This study developed a 2D turn-based RPG in where a designer defines a game performance curve and the DDA mechanism adjusts the game to align the player to that path. The player controls four characters in a turn-based spell battle game with four different enemy types. There are three different enemy types with strengths and weaknesses to battle. The DDA system struggled to perform

			when confronted with unpredictable player behaviour.

References

Abdessalem, H. B., & Frasson, C. (2017). Real-time brain assessment for adaptive virtual reality game: a neurofeedback approach. In *International Conference on Brain Function Assessment in Learning* (pp. 133-143): Springer.

Adams, E. (2014). *Fundamentals of game design*: Pearson Education.

Afergan, D., Peck, E. M., Solovey, E. T., Jenkins, A., Hincks, S. W., Brown, E. T., Chang, R., & Jacob, R. J. K. (2014). Dynamic difficulty using brain metrics of workload. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14* (pp. 3797-3806).

Al-Elq, A. H. (2010). Simulation-based medical teaching and learning. *Journal of family and Community Medicine, 17*(1), 35.

Aldekhyl, S., Cavalcanti, R. B., & Naismith, L. M. (2018). Cognitive load predicts point-of-care ultrasound simulator performance. *Perspectives on medical education, 7*(1), 23-32.

Alexander, J. T., Sear, J., & Oikonomou, A. (2013). An investigation of the effects of game difficulty on player enjoyment. *Entertainment Computing, 4*(1), 53-62.

Alves, T., Gama, S., & Melo, F. S. (2018). Flow adaptation in serious games for health. In *2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH)* (pp. 1-8): IEEE.

Andrade, K. d. O., Pasqual, T. B., Caurin, G. A., & Crocomo, M. K. (2016). Dynamic difficulty adjustment with Evolutionary Algorithm in games for rehabilitation robotics. In *2016 IEEE International Conference on Serious Games and Applications for Health (SeGAH)* (pp. 1-8): IEEE.

Andrew, Tjokrosetio, A. N., & Chowanda, A. (2020). Dynamic difficulty adjustment with facial expression recognition for improving player satisfaction in a survival horror game. *ICIC Express Letters, 14*(11), 1097-1104.

Avi Shena, B. S., Sitohang, B., & Rukmono, S. A. (2019). Application of Dynamic Difficulty Adjustment on Evidence-centered Design Framework for Game Based Learning. In *2019 International Conference on Data and Software Engineering (ICoDSE)* (pp. 1-6): IEEE.

Baldwin, A., Johnson, D., & Wyeth, P. A. (2014). The effect of multiplayer dynamic difficulty adjustment on the player experience of video games. In *Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing systems - CHI EA '14* (pp. 1489-1494).

Bian, D., Wade, J., Swanson, A., Weitlauf, A., Warren, Z., & Sarkar, N. (2019). Design of a physiology-based adaptive virtual reality driving platform for individuals with ASD. *ACM Transactions on Accessible Computing, 12*(1).

Bontchev, B. (2016). Adaptation in Affective Video Games: A Literature Review. *Cybernetics and Information Technologies, 16*(3), 3-34.

Bontchev, B., & Georgieva, O. (2018). Playing style recognition through an adaptive video game. *Computers in Human Behavior, 82*, 136-147.

Burns, A., & Tulip, J. (2017). Detecting flow in games using facial expressions. In *2017 IEEE Conference on Computational Intelligence and Games (CIG)* (pp. 45-52): IEEE.

Catarino, J., & Martinho, C. (2019). Procedural progression model for smash time. In *IEEE Conference on Computatonal Intelligence and Games, CIG* (Vol. 2019-August).

Chanel, G., & Lopes, P. (2020). User Evaluation of Affective Dynamic Difficulty Adjustment Based on Physiological Deep Learning. In *International Conference on Human-Computer Interaction* (Vol. 12196 LNAI, pp. 3-23): Springer.

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education, 59*(2), 661-686.

Conte, A., Scarsini, M., & Sürücü, O. (2016). The impact of time limitation: Insights from a queueing experiment. *Judgment and Decision Making, 11*(3), 260-274.

Csikszentmihalyi, M., Abuhamdeh, S., & Nakamura, J. (2014). Flow. In *Flow and the foundations of positive psychology* (pp. 227-238): Springer.

De Freitas, S. (2018). Are games effective learning tools? A review of educational games. *Journal of Educational Technology & Society, 21*(2), 74-84.

De la Cruz, S. M., Portocarrero, S. S., & Shiguihara-Juárez, P. (2018). Adaptive Serious Game as a Learning Approach for Microbiology. In *2018 Congreso Internacional de Innovación y Tendencias en Ingeniería (CONIITI)* (pp. 1-6): IEEE.

Demediuk, S., Tamassia, M., Li, X., & Raffe, W. L. (2019). Challenging AI. In *Proceedings of the Australasian Computer Science Week Multiconference on - ACSW 2019* (pp. 1-7).

Denisova, A., & Cairns, P. (2015). Adaptation in digital games: The effect of challenge adjustment on player performance and experience. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play - CHI PLAY '15* (pp. 97-101).

Dideriksen, A., Reuter, C., Patry, T., Schnell, T., Hoke, J., & Faubert, J. (2018). Define Expert - Characterizing Proficiency for Physiological Measures of Cognitive Workload. In *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2018* (2018 ed.). I/ITSEC Knowledge Repository.

El-Habr, C., Garcia, X., Paliyawan, P., & Thawonmas, R. (2019). Runner: A 2D platform game for physical health promotion. *Softwarex, 10*.

Ewing, K. C., Fairclough, S. H., & Gilleade, K. (2016). Evaluation of an Adaptive Game that Uses EEG Measures Validated during the Design Process as Inputs to a Biocybernetic Loop. *Frontiers in Human Neuroscience, 10*, 223.

Fernandes, W. R., & Levieux, G. (2019). delta-logit : Dynamic Difficulty Adjustment Using Few Data Points. In *Joint International Conference on Entertainment Computing and Serious Games* (Vol. 11863, pp. 158-171): Springer.

Galy, E., Cariou, M., & Mélan, C. (2012). What is the relationship between mental workload factors and cognitive load types? *International Journal of Psychophysiology, 83*(3), 269-275.

Gu, X., & Blackmore, K. (2017). Characterisation of academic journals in the digital age. *Scientometrics, 110*(3), 1333-1350.

Haji, F. A., Rojas, D., Childs, R., de Ribaupierre, S., & Dubrowski, A. (2015). Measuring cognitive load: performance, mental effort and simulation task complexity. *Medical education, 49*(8), 815-827.

Hamari, J., Shernoff, D. J., Rowe, E., Coller, B., Asbell-Clarke, J., & Edwards, T. (2016). Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. *Computers in Human Behavior, 54*, 170-179.

Hendrix, M., Bellamy-Wood, T., McKay, S., Bloom, V., & Dunwell, I. (2019). Implementing adaptive game difficulty balancing in serious games. *IEEE Transactions on Games, 11*(4), 320-327.

Hooshyar, D., Malva, L., Yang, Y., Pedaste, M., Wang, M., & Lim, H. (2021). An adaptive educational computer game: Effects on students' knowledge and learning attitude in computational thinking. *Computers in Human Behavior, 114*.

Hunicke, R. (2005). The case for dynamic difficulty adjustment in games. In *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology* (pp. 429-433): ACM.

Hunicke, R., & Chapman, V. (2004). AI for Dynamic Difficulty Adjustment in Games. *Proceedings of the Challenges in Game Artificial Intelligence AAAI Workshop*, 91–96.

Ilici, L., Wang, J., Missura, O., & Gärtner, T. (2012). Dynamic difficulty for checkers and Chinese chess. In *2012 IEEE Conference on Computational Intelligence and Games (CIG)* (pp. 55-62): IEEE.

Kapp, K. M. (2012). *The gamification of learning and instruction: game-based methods and strategies for training and education*: John Wiley & Sons.

Kelley, C. R. (1969). What is adaptive training? *Human Factors, 11*(6), 547-556.

Kim, B. (2001). Social constructivism. *Emerging perspectives on learning, teaching, and technology, 1*(1), 16.

Kozulin, A., Ageyev, V. S., Gindis, B., & Miller, S. M. (2003). *Vygotsky's educational theory in cultural context*: Cambridge University Press.

Landsberg, C. R., Van Buskirk, W. L., Astwood Jr, R. S., Mercado, A. D., & Aakre, A. J. (2010). Adaptive training considerations for use in simulation-based systems. In: NAVAL AIR WARFARE CENTER TRAINING SYSTEMS DIV ORLANDO FL.

Li, Y.-N., Yao, C., Li, D.-J., & Zhang, K. (2014). Adaptive difficulty scales for Parkour games. *Journal of Visual Languages & Computing, 25*(6), 868-878.

Lim, T., Louchart, S., Suttie, N., Hauge, J. B., Stanescu, I. A., Ortiz, I. M., Moreno-Ger, P., Bellotti, F., Carvalho, M. B., & Earp, J. (2014). Narrative serious game mechanics (NSGM)–insights into the narrative-pedagogical mechanism. In *International conference on serious games* (pp. 23-34): Springer.

Liu, C., Agrawal, P., Sarkar, N., & Chen, S. (2009). Dynamic Difficulty Adjustment in Computer Games Through Real-Time Anxiety-Based Affective Feedback. *International Journal of Human-Computer Interaction, 25*(6), 506-529.

Lora Ariza, D. S., Sánchez-Ruiz, A. A., & González-Calero, P. A. (2019). Towards Finding Flow in Tetris. In *International Conference on Case-Based Reasoning* (Vol. 11680 LNAI, pp. 266-280): Springer.

Mainieri, B. O., Azevedo, V. L., Braga, P. H. C., & Omar, N. (2018). Development and Assessment of an Adaptive Difficulty Arithmetic Game Based Learning Object. In *2018 XIII Latin American Conference on Learning Technologies (LACLO)* (pp. 232-239): IEEEE.

Masanobu, E., B, H. D. F., & Mikami, K. (2017). Dynamic Pressure Cycle Control: Dynamic Diffculty Adjustment beyond the Flow Zone. In *2017 Nicograph International (NicoInt)* (pp. 9-14).

Mayefsky, E., Anene, F., & Sirota, M. (2003). Algorithms - Minimax. In.

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic reviews, 4*(1), 1.

Moon, H.-S., & Seo, J. (2020). Dynamic Difficulty Adjustment via Fast User Adaptation. In *Adjunct Publication of the 33rd Annual ACM Symposium on User Interface Software and Technology* (pp. 13–15). Virtual Event, USA: Association for Computing Machinery.

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational psychology review, 19*(3), 309-326.

Nagle, A., Novak, D., Wolf, P., & Riener, R. (2014). The effect of different difficulty adaptation strategies on enjoyment and performance in a serious game for memory training. In *2014 IEEE 3nd International Conference on Serious Games and Applications for Health (SeGAH)* (pp. 1-8): IEEE.

Nagle, A., Novak, D., Wolf, P., & Riener, R. (2015). Increased enjoyment using a tablet-based serious game with regularly changing visual elements: A pilot study. *Gerontechnology, 14*(1).

Negini, F., Mandryk, R. L., & Stanley, K. G. (2014). Using affective state to adapt characters, NPCs, and the environment in a first-person shooter game. In *2014 IEEE Games Media Entertainment* (pp. 1-8): IEEE.

Ninaus, M., Tsarava, K., & Moeller, K. (2019). A pilot study on the feasibility of dynamic difficulty adjustment in game-based learning using heart-rate. In *International Conference on Games and Learning Alliance* (Vol. 11899 LNCS, pp. 117-128): Springer.

Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional science, 32*(1/2), 1-8.

Pagalyte, E., Mancini, M., & Climent, L. (2020). Go with the Flow: Reinforcement Learning in Turn-based Battle Video Games. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents* (pp. Article 44). Virtual Event, Scotland, UK: Association for Computing Machinery.

Papadimitriou, S., Chrysafiadi, K., & Virvou, M. (2019). FuzzEG: Fuzzy logic for adaptive scenarios in an educational adventure game. *Multimedia Tools and Applications, 78*(22), 32023-32053.

Park, S., Sim, H., & Lee, W. (2014). Dynamic Game Difficulty Control by Using EEG-based Emotion Recognition. *International Journal of Control and Automation, 7*(3), 267-272.

Parnandi, A., & Gutierrez-Osuna, R. (2014). A comparative study of game mechanics and control laws for an adaptive physiological game. *Journal on Multimodal User Interfaces, 9*(1), 31-42.

Pato, V. M. Á., & Delgado-Mata, C. (2013). Dynamic Difficulty Adjusting Strategy for a Two-player Video Game. *Procedia Technology, 7*, 315-321.

Peirce, N., Conlan, O., & Wade, V. (2008). Adaptive Educational Games: Providing Non-invasive Personalised Learning Experiences. In *2008 Second IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning* (pp. 28-35).

Pfau, J., Smeddinck, J. D., & Malaka, R. (2020). Enemy Within: Long-term Motivation Effects of Deep Player Behavior Models for Dynamic Difficulty Adjustment. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-10).

Plass, J. L., Homer, B. D., Pawar, S., Brenner, C., & MacNamara, A. P. (2019). The effect of adaptive difficulty adjustment on the effectiveness of a game to develop executive function skills for learners of different ages. *Cognitive Development, 49*, 56-67.

Popescu, M.-M., Roceanu, I., Earp, J., Ott, M., & Ger, P. M. (2012). ASPECTS OF SERIOUS GAMES CURRICULUM INTEGRATION-A TWO-FOLDED APPROACH. *eLearning & Software for Education*(2).

Prendinger, H., Puntumapon, K., & Madruga, M. (2014). Extending Real-Time Challenge Balancing to Multiplayer Games: A Study on Eco-Driving. *IEEE Transactions on Computational Intelligence and AI in Games, 8*(1), 27-32.

Rauterberg, M. (1995). About a framework for information and information processing of learning systems. In *Information System Concepts* (pp. 54-69): Springer.

Ravyse, W. S., Seugnet Blignaut, A., Leendertz, V., & Woolner, A. (2016). Success factors for serious games to enhance learning: a systematic review. *Virtual Reality, 21*(1), 31-58.

Reidy, L., Chan, D., Nduka, C., & Gunes, H. (2020). Facial Electromyography-based Adaptive Virtual Reality Gaming for Cognitive Training. In *Proceedings of the 2020 International Conference on Multimodal Interaction* (pp. 174–183). Virtual Event, Netherlands: Association for Computing Machinery.

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *The Journal of the Learning sciences, 13*(3), 273-304.

Rietveld, A., Bakkes, S., & Roijers, D. (2014). Circuit-adaptive challenge balancing in racing games. In *2014 IEEE Games Media Entertainment* (pp. 1-8): IEEE.

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*: Guilford Publications.

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion, 30*(4), 344-360.

Salas, E., Rosen, M. A., Held, J. D., & Weissmuller, J. J. (2009). Performance measurement in simulation-based training: A review and best practices. *Simulation & gaming, 40*(3), 328-376.

Sampayo-Vargas, S., Cope, C. J., He, Z., & Byrne, G. J. (2013). The effectiveness of adaptive difficulty adjustments on students' motivation and learning in an educational computer game. *Computers & Education, 69*, 452-462.

Schwabe, L., & Wolf, O. T. (2010). Learning under stress impairs memory formation. *Neurobiology of learning and memory, 93*(2), 183-188.

Seyderhelm, A. J., Blackmore, K. L., & Nesbitt, K. (2019). Towards Cognitive Adaptive Serious Games: A Conceptual Framework. In *Joint International Conference on Entertainment Computing and Serious Games* (pp. 331-338): Springer.

Sharek, D., & Wiebe, E. (2014). Measuring video game engagement through the cognitive and affective dimensions. *Simulation & gaming, 45*(4-5), 569-592.

Silva, M. P., Silva, V. d. N., & Chaimowicz, L. (2017). Dynamic difficulty adjustment on MOBA games. *Entertainment Computing, 18*, 103-123.

Sitzmann, T. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. *Personnel psychology, 64*(2), 489-528.

Spiel, K., Bertel, S., & Kayali, F. (2019). Adapting gameplay to eye movements – An exploration with Tetris. In *CHI PLAY 2019 - Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play* (pp. 687-695).

Stein, A., Yotam, Y., Puzis, R., Shani, G., & Taieb-Maimon, M. (2018). EEG-triggered dynamic difficulty adjustment for multiplayer games. *Entertainment Computing, 25*, 14-25.

Suaza, J., Gamboa, E., & Trujillo, M. (2019). A Health Point-Based Dynamic Difficulty Adjustment Strategy for Video Games. In *Joint International Conference on Entertainment Computing and Serious Games* (Vol. 11863 LNCS, pp. 436-440): Springer.

Sweller, J. (2011). Cognitive load theory. In *Psychology of learning and motivation* (Vol. 55, pp. 37-76): Elsevier.

Tadayon, R., Amresh, A., McDaniel, T., & Panchanathan, S. (2018). Real-time stealth intervention for motor learning using player flow-state. In *2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH)* (pp. 1-8): IEEE.

Tahai, L., Wallace, J. R., Eckhardt, C., & Pietroszek, K. (2019). Scalebridge: Design and evaluation of adaptive difficulty proportional reasoning game for children. In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)* (pp. 1-4): IEEE.

Tan, K. L., Tan, C. H., Tan, K. C., & Tay, A. (2009). Adaptive game AI for Gomoku. In *2009 4th International Conference on Autonomous Robots and Agents* (pp. 507-512): IEEE.

Tang, Y., Liang, J., Hare, R., & Wang, F. Y. (2020). A Personalized Learning System for Parallel Intelligent Education. *IEEE Transactions on Computational Social Systems, 7*(2), 352-361.

Tomlinson, S., & Melder, N. (2013). An Architecture Overview for AI in Racing Games. *Game AI Pro: Collected Wisdom of Game AI Professionals*, 471.

Van Lankveld, G., Spronck, P., Van Den Herik, H. J., & Rauterberg, M. (2009). Incongruity-based adaptive game balancing. In *Advances in computer games* (pp. 208-220): Springer.

van Oostendorp, H., van der Spek, E. D., & Linssen, J. (2014). Adapting the Complexity Level of a Serious Game to the Proficiency of Players. *EAI Endorsed Transactions on Game-Based Learning, 1*(2).

Vicencio-Moreira, R., Mandryk, R. L., & Gutwin, C. (2014). Balancing multiplayer first-person shooter games using aiming assistance. In *2014 IEEE Games Media Entertainment* (pp. 1-8): IEEE.

Vlachopoulos, D., & Makri, A. (2017). The effect of games and simulations on higher education: a systematic literature review. *International Journal of Educational Technology in Higher Education, 14*(1), 22.

Woo, J.-C. (2014). Digital game-based learning supports student motivation, cognitive success, and performance outcomes. *Journal of Educational Technology & Society, 17*(3), 291-307.

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, *105*(2), 249-265.

Xiang, N., Yang, L., & Zhang, M. (2013). Dynamic Difficulty Adjustment by Facial Expression. In *Informatics and Management Science V* (pp. 761-768).

Xue, S., Wu, M., Kolen, J., Aghdaie, N., & Zaman, K. A. (2017). Dynamic difficulty adjustment for maximized engagement in digital games. In *Proceedings of the 26th International Conference on World Wide Web Companion* (pp. 465-471): International World Wide Web Conferences Steering Committee.

Yun, C., Shastri, D., Pavlidis, I., & Deng, Z. (2009). O' game, can you feel my frustration?: Improving user's gaming experience via stresscam. In *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*.

Yun, C., Trevino, P., Holtkamp, W., & Deng, Z. (2010). PADS: enhancing gaming experience using profile-based adaptive difficulty system. In *Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games* (pp. 31-36): ACM.

Zohaib, M. (2018). Dynamic Difficulty Adjustment (DDA) in Computer Games: A Review. *Advances in Human-Computer Interaction, 2018*, 1-12.

Zook, A., & Riedl, M. O. (2014). Temporal game challenge tailoring. *IEEE Transactions on Computational Intelligence and AI in Games, 7*(4), 336-346.

View publication stats