Measuring Perceived Challenge in Digital Games: Development & Validation of the Challenge Originating from Recent Gameplay Interaction Scale (CORGIS)

Alena Denisova*

City, University of London, London, EC1V 0HB, UK

Paul Cairns

University of York, York, YO10 5GH, UK

Christian Guckelsberger

Queen Mary University of London, London, E1 4NS, UK

David Zendle

York St John University, York, YO31 7EX, UK

Abstract

Challenge is a key element of digital games, but a clear conceptualisation and operationalisation of this player experience were long missing. This made it hard for game developers to design for a well-balanced experience across different skill-sets, and impeded the synthesis of challenge-related games research. To overcome this, we introduce a systematic, extensive, and reliable instrument to evaluate the level of players' perceived challenge in digital games. We conceptualise challenge based on a survey of related literature in games user research, design and AI, as well as interviews with researchers and players. Exploratory factor analysis (N=394) highlights four components of experienced challenge: performative, emotional, cognitive and decision-making challenge. Refinement of the items allowed us to devise the *Challenge Originating from Recent Gameplay Interaction Scale* (CORGIS), which has been further validated in a study with nearly 1,000 players. The question-

*Corresponding author

Email address: alena.denisova@city.ac.uk (Alena Denisova)

Preprint submitted to International Journal of Human Computer Studies August 3, 2019

naire exhibits good construct validity for use by both game developers and researchers to quantify players' challenge experiences.

Keywords: perceived challenge, difficulty, questionnaire, player experience, digital games

1. Introduction

> *"Play can be pleasurable when it hurts, offends, challenges us and teases us, and even when we are not playing."*
> — Miguel Sicart, *Play Matters*, 2014

Challenge is a central constituent of the gameplay in the majority of digital games (Feil and Scattergood, 2005), and is widely believed to play a crucial role in making games enjoyable (Vorderer et al., 2003). *Play*, in the more general sense, is often understood as "free movement within a more rigid structure" (Salen and Zimmerman, 2004). Unlike application-system users who expect to complete a task in the most efficient and unobstructed way, game players enjoy being challenged in the mastery of a game, preferably in various ways (Davis et al., 2005; Shneiderman, 1982).

A player's perceived challenge is connected to many other experiences. For instance, a level of challenge that is neither too hard nor too easy is essential to achieving the state of *flow* (Csikszentmihalyi, 1990) and contributes to players experiencing *immersion* (Ermi and Mäyrä, 2005; Jennett et al., 2008). However, it is hard to design such optimal challenge for players with varying levels of skill, experience, and motivation. Despite the evident importance of challenge for games, little research has gone into exploring exactly how challenge is perceived by players. Both in industry and research, challenge is presently assessed ultimately through reports of playtesters, potentially complementing automated playtesting with bots (Nielsen et al., 2015). These individual reports, however, usually only capture a small part of what constitutes the experience of challenge and are typically unstructured, rendering a comparison across different games or players infeasible.

Though there have been efforts in recent years to better define challenge (Cole et al., 2015; Denisova et al., 2017; Bopp et al., 2018), there is currently no robust operationalisation and measure of challenge in video games.

Based on a systematic review of nine widely used and easily accessible questionnaires and an analysis of their challenge-related items, Denisova et al. (2017) have identified a lack of depth, fragmented items, and, in some cases, the absence of statistical validation. The main drawback of the existing tools, however, is that they were neither designed to measure the *experience*, i.e. the player's subjective *perception* of challenge, nor do they measure and quantify all of its facets, such as emotional and social challenge.

To assess the complex challenge structures of modern games, we need a questionnaire comprehensive enough to capture experiences of players with varied skill and experience playing different kinds of games. A more reliable instrument would allow game developers to assess a larger spectrum of challenge types more precisely and in a comparable way. At the same time, researchers would get a better tool to both eliminate challenge as confounding variable and to scientifically explore challenge in games further in standardised experiments. Together, this would open up possibilities for more diversified and richer game-playing experiences.

The primary objective of this research was to create and validate a questionnaire to measure challenge in video games which overcomes the present limitations. This paper documents the resulting process of conceptualising the construct of perceived challenge in video games based on the analysis of relevant literature and interviews with players and games researchers. It furthermore summarises the development and validation of the *Challenge Originating from Recent Gameplay Interaction Scale* (CORGIS) through two large-scale surveys. The result is a robust, empirically supported measure of perceived challenge, which has the potential to inform, unify and thus advance research into player experience around challenge in games.

Structure of this Article. Our questionnaire development process in Figure 1 followed the procedures outlined by Kline (2014). It involved the following steps, which are reflected in the structure of this article:

1. **Reviewing related work** to operationalise challenge as player experience, to identify the key factors that might lead to challenging experiences and to explore the relationships between challenge and other player experiences. We elaborate on this in Section 2.

2. **Conducting semi-structured interviews** with video game players and researchers to fill in any gaps identified in the existing literature. Section 3 provides a detailed discussion of the interviews.

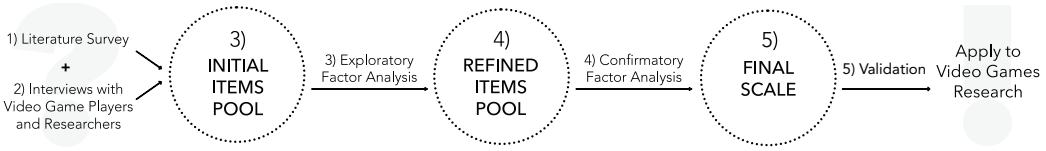

Figure 1: The stages of the questionnaire development process of CORGIS.

3. **Generating an initial items pool** in Section 4, based on the themes identified in the literature review and interviews. We collect data from a range of players about their experiences of challenge. Our approach to data analysis is documented in Section 5. We perform *Exploratory Factor Analysis* (EFA) to assess the factor structure and internal consistency of the larger item pool, documented in Section 6.

4. **Refining the items pool** based on the EFA. We gather data from another sample of players about their experiences of challenge. We perform *Confirmatory Factor Analysis* (CFA) to assess the construct validity of the reduced item pool. Section 7 provides the results of the CFA.

5. **Validating the questionnaire** using datasets from three different games with the distinct prevalence of specific challenge types. Section 8 provides the validation of the CORGIS using these three games.

We discuss our contributions and the limitations of the work (Sec. 9), followed by a conclusion and future work (Sec. 10).

2. Perceived Challenge in Video Games

Gameplay characterises the core activity of a game, a series of actions performed by the player and game actors as well as their associated feedback or outcomes (Vorderer et al., 2003). Challenge and gameplay are closely intertwined, or as Adams (2014) notes:

> *"Gameplay is challenges and actions that entertain. People enjoy a challenge, as long as they can reasonably expect to accomplish it. People also try a challenge they do not expect to meet if the risk is low and the reward is high. Challenges create tension and drama. At the simplest level, presenting players with a challenge amounts to asking a question: "Can you do it?" They'll enjoy trying to prove that you can."*

The notions of *'difficulty'*/*'difficult'* and *'challenge'*/*'challenging'* are often used interchangeably in the literature (cf. e.g. Malone, 1982). To understand the concept of *perceived challenge*, we must first clarify the differences between these two notions and their intrinsic and relational forms. Firstly, *'difficulties'* and *'challenges'* are determined by the objective of the game and the barriers that prevent the player from achieving it. They denote a task or problem, rather than specific experience. Lomas et al. (2017) define *'difficulty'* as "the probability of task failure", which can be objectively measured in the game. In contrast, *'difficult'* and *'challenging'* as relational attributes refer to how difficulties and challenges are experienced by the player.

Secondly, *'difficult'* and *'challenging'* are commonly used with different valence. We use the word *'difficult'* for tasks that we struggle to continue with, but *'challenging'* is used more positively to describe a demanding, yet stimulating task or problem. A difficult game can be frustrating and cause discomfort. A challenging game is stimulating, and challenged players are motivated to respond to tasks, their actions make a difference, and they feel in control over the outcomes of these actions (Lazzaro, 2004).

When measuring *perceived challenge*, we want to assess how a specific player experiences both the *challenges* and *difficulties* posed by a particular game, i.e. how *challenging* and *difficult* they find the game. Despite saying 'perceived *challenge*', we thus want to capture both positive and negative valences. In the remaining text though, we use the notions of 'difficult' and 'challenging' free of valence if not stated differently. Crucially, perceived challenge depends on the player's skills, which are shaped by preceding gameplay experiences. Given this temporal dependency, we refine the formulation of our goal to the assessment of challenge *originating from recent gameplay*.

2.1. Types of Perceived Challenge in Games

Challenge is a multi-faceted experience. In order to develop an instrument capable of measuring the experience of challenge in its full breadth, we need to differentiate between different types of challenge.

Physical (performative) challenge addresses a player's abilities with regards to the speed and reaction times, as well as precision and accuracy of performing actions (Cox et al., 2012). Certain video games also exercise the player's physical endurance, dexterity, and coordination (Schell, 2014). Physical challenges were dominant in early arcade games, e.g. *Space Invaders*, but are also present in modern games such as *Overwatch*, which require good reaction and mastery of controls.

Performative Challenge: Challenge that addresses the player's physical limitations to interact with the game, i.e. the speed and accuracy with which actions can be performed.

Cognitive challenge, conversely, refers to a player's memory, observation, and problem-solving capacities. Typically, cognitive challenges require players to have good spatial and logical reasoning, decision-making, and planning to make progress in the game, and are often found in puzzle games, e.g. *Candy Crush*, or in strategy games, such as *Civilization*. The player has to invest cognitive effort to predict the consequences of their actions or comprehend ambiguous elements of the narrative or storyline.

A less prominent challenge type is social challenge. It can be viewed as a sub-type of cognitive challenge (Malone, 1981), in that it requires the player to deal with hidden information. It arises during play against or with human or AI players and relates to the capacity to read an opponent, predicting their moves and making split-second decisions that can aid players in deceiving opponents (Denisova et al., 2017). Conversely, success as a team depends on being able to predict the targets of teammates and to coordinate joint actions (Bopp et al., 2018). Beyond competitive action titles, games such as *L.A. Noire* provide social challenges, in that they require the player to read and disambiguate social cues to prevent deceit.

Cognitive Challenge: Challenge that addresses the player's cognitive and problem-solving capacities. The player has to invest cognitive effort to predict the consequences of actions or comprehend ambiguous elements of the narrative or the storyline.

Emotional challenge represents a third, less prominent but yet important type of challenge first introduced by Cole et al. (2015). Games stimulate emotional challenge by providing players with a compelling narrative or story, through ambiguous or difficult material or by using strong characters for which the player can feel empathy. A player can overcome emotional challenge by resolving tension in the narrative and by identifying with characters (Cole et al., 2015). While games like *Journey* or *Life is Strange* do not require players to master controls particularly well in order to advance in the game, they establish an emotional connection with the game world and allow players to relate and feel connected to the characters and the story. An

emotional connection can also emerge when players take on responsibility for others as in *Papers, Please*; when experiencing betrayal; or when losing valued or trustworthy companions such as in escorting missions (Isbister, 2016). Nonetheless, players can also be emotionally challenged under time constraints (Isbister, 2016), by struggling to do all the things they want to do. In their survey of emotionally challenging experiences in games, Bopp et al. (2018) complement physical, cognitive and social challenge with the three additional challenge types of *decisions & actions*, *difficult themes* and *intense emotions*, which we here subsume under emotional challenge.

> **Emotional Challenge:** Challenge which confronts the player with emotionally salient material or the use of strong characters, and a captivating story. A player cannot overcome emotional challenge with skill or dexterity, but by resolving tension in the narrative, by assessing their identification with game characters, and by resolving ambiguities.

Different game genres come with varying types of challenges, but the boundaries are blurred and modern digital games usually expose players to several challenge types at once. The role-playing games (RPGs) of the *Shin Megami Tensei* series, for instance, focus mainly on mental challenges within battles but also creates emotional challenge in the form of moral dilemmas. Similarly, quests in *The Elder Scrolls V: Skyrim* not only require good performative skills for battle but also mental capacities for solving puzzles.

The complexity in the challenge structure of games is emphasised in the struggle of researchers to develop artificial intelligence (AI) agents that play equally well in a wide range of games (Bontrager et al., 2016). Our instrument for measuring challenge must, therefore, be comprehensive enough to capture perceived challenge in any digital game, while also being sufficiently specific to differentiate between perceived challenge types that potentially lead to different experiences – the emotional involvement in the classic, performance-oriented *Winter Games* would certainly differ from the emotions triggered in the slow-paced *Journey*.

2.2. Perceived Challenge as Player Experience

Different types of challenge can invoke different player experiences. In this section, we identify both (i) the factors that influence perceived challenge, and, vice versa, (ii) other player experiences that arise from being challenged.

Our review rests on an in-depth analysis of related work in games user research, game design and game AI.

Difficulty $\times$ *Skill* $\rightarrow$ *Perceived Challenge.* A player's perception of challenge largely depends on the difficulty of the game (Adams, 2014; Malone, 1981, 1982). This has been empirically demonstrated by Yannakakis and Hallam (2007, 2009), who managed to improve a player's 'entertainment value' by letting online mechanism adapt game parameters that control perceived challenge. A difficult game is perceived as hard to play if the player lacks the relevant skills or expertise. The ability to successfully face challenges thus influences the players' perception of difficulty; a more skilled player would experience less challenge when overcoming the same obstacle than a player with fewer experience and relevant skills. Thus, the balance of player skill and game difficulty represent one of the key determinants of performative and cognitive challenge.

As the difficulty in the game increases with time, players should be able to learn and improve their skills in order to eventually master the game. If the game supports the improvement of one's abilities and skills, players should eventually experience optimal challenge, i.e. a perfect match between the game difficulty and the players' skill level. This can be complemented and even replaced with a *dynamic difficulty adjustment* (Hunicke, 2005) that tries to adapt the difficulty of the game based on the player's performance 'in real time'. One example is the 'rubber-band mechanics' in the *Mario Kart* series, through which players that fall behind in the race receive more powerful and rare items than those in front.

Uncertainty $\rightarrow$ *Perceived Challenge.* Perceived challenge can also emerge as a result of players feeling uncertain (Malone, 1981). Being unable to predict the outcome of one's actions, failing to read the opponent, or feeling uncertain about the best possible tactic for a battle increases one's perception of challenge. Being unable to predict whether the player will succeed or fail has been claimed to provide a strong motivation to play (Loftus and Loftus, 1983) and 'close call' games are played for longer (Lomas et al., 2017). Similarly, other empirical studies show that outcome uncertainty relates to the feeling of suspense, which in turn increases enjoyment (Abuhamdeh et al., 2015). This also represents one explanation for the appeal of competition (Poels et al., 2007; Vorderer et al., 2003).

Playing task-based or competitive games is, however, enjoyable as long as the outcome of the task remains uncertain and the balance between challenge

and mastery is achieved (Klimmt et al., 2007). Players often experience suspense and curiosity as a result of feeling uncertain with regards to being able to cope with challenges as they arise (Klimmt, 2003).

Success + Failure → Perceived Challenge. Challenge can be evaluated through players' perception of their performance, dependent on their experiences of success or failure. This is by no means limited to performative and cognitive challenge; Bopp et al. (2018) have identified 'achievement' as a key element of emotionally challenging experiences. Similarly, AI researchers have assessed the variety and depth of challenge in different games by comparing the relative performance of AI agents of various complexity. This is based on the assumption that in games with more varied challenge, we can expect a larger score difference between a simple and a sophisticated AI (Nielsen et al., 2015).

However, repeated failure in the game can lead to frustration, especially if it remains explainable. Conversely, if the game provides sufficient feedback to the player, failure can be seen as an essential part of learning. Without failing, players would quickly deem a game boring (Juul, 2009). Failure not only makes winning more enjoyable, but it also makes players readjust their perception of a game: *"Failure adds content by making the player see new nuances in a game."* (Juul, 2009). According to Rouse III (2010): *"Players need to blame only themselves for not succeeding, but at the same time the game must be challenging enough so that they do not succeed right away."*

Perceived Challenge → Flow + Immersion. Perceived challenge can give rise to other player experiences or can be considered a mediating variable between the interaction of a game's difficulty and the player's skill to other player experiences. In particular, optimal challenge relative to a player's skills is a key requirement for reaching the state of *flow* (Csikszentmihalyi, 1990; Chen, 2007) – an optimal experience evoked by high levels of engagement in an activity. In this state, a person experiences high concentration as they become fully absorbed by their play. Flow is the key target experience for dynamic difficulty adjustment algorithms in games.

Flow, however, is not the only experience that emerges as a result of well-balanced challenge. Players feel more *immersed* when the balance between the game difficulty and their skill level is matched fairly closely (Cox et al., 2012; Jennett et al., 2008). Ermi and Mäyrä (2005) distinguish challenge-based immersion as a unique experience and hypothesise that it is affected

by both the challenge of 'pace' (i.e. performative challenge) and 'cognitive challenge'. Cox et al. (2012) report from experiments that increased physical demand does not per se increase immersion, but adding time pressure increases performative and cognitive challenge and, consequently, immersion.

Perceived Challenge $\rightarrow$ *Competence.* Optimal challenge is also linked to the sense of competence (Abuhamdeh and Csikszentmihalyi, 2012; Ryan et al., 2006) – the experience arising from one's ability to meet the requirements of tasks they have or want to complete. Deci and Ryan (2000) state: "It is success at optimally challenging tasks that allows people to feel a true sense of competence.". Perceived competence represents a key component in the *Self-Determination Theory* (SDT) of motivation which, applied to videogames, seeks to explain people's ongoing engagement with and enjoyment of a game.

Perceived Challenge $\leftrightarrow$ *Affect.* As part of their research on SDT, Ryan et al. (2006) also highlight that optimal challenge and competence relate to a player's emotional responses to a video game. Players might experience positive or negative emotional responses not only from emotional challenge. When a challenge is beyond one's abilities, this may lead to anxiety, or if the player does not feel challenged enough – to boredom (Csikszentmihalyi, 1990). Being optimally challenged leads players to experience enjoyment (Abuhamdeh and Csikszentmihalyi, 2012) and pleasure (Hunicke et al., 2004). If the player cannot achieve this experience of optimal challenge, this can lead to tension, which for more experienced players can turn into irritation, disappointment, anger, and frustration. Frustration and irritation are particular emotions that emerge from a mismatch between difficulty and skills (Poels et al., 2007).

High levels of perceived challenge, however, do not always lead to negative experiences. Petralito et al. (2017) demonstrated that some players value excessive difficulty, as in the *Meat Boy* series, which is linked to their experienced sense of achievement and learning moments in the game. Moreover, negative emotions experienced in the game are not always perceived as undesirable by the players: in-game loss, character attachment and the lack of agency can be viewed as emotionally rewarding and thought-provoking experiences (Bopp et al., 2016). And this, in turn, can lead to high perceived emotional challenge and thus a bidirectional relationship between perceived challenge and affect.

When considering changes in the game difficulty throughout the whole experience of playing a specific game, the relationship between enjoyment

and perceived challenge usually becomes even more complex. Klimmt et al. (2009) show that players tend to prefer lower levels of difficulty at the start of the game; as players make progress though, they might enjoy it more when the difficulty offered by the game is slightly higher than their perceived level of skills and challenge (Bateman et al., 2011; Klarkowski et al., 2016).

> **Perceived Challenge:** A player experience that arises from one's interaction with a game's intrinsic challenges at a particular skill level.

The proposed relationships between perceived challenge and other player experiences are mostly based on fragmented literature and these links are not well supported by evidence. This is largely due to the lack of a systematic approach to defining and assessing perceived challenge. The review also highlights that current research into perceived challenge focuses primarily on performative and cognitive challenge, while emotional challenge has only received little attention. To be able to verify these links and exploit them in our questionnaire, we have conducted a semi-structured qualitative study reported in the next section.

3. A Semi-Structured Qualitative Study for Item Development

The reviewed literature gives a good indication of the variety of challenges that games are able to present to players. However, little work has been done to explore how these challenges are perceived from the player perspective. While challenges may have diverse roots, they may lead to a similar felt experience or, conversely, similar game features may lead to differently perceived challenges because of player individual differences and contextual factors (Denisova and Cairns, 2015).

To explore players' experience of challenge more flexibly and in depth, we conducted a semi-structured qualitative study (Blandford, 2013). In particular, we were looking for the attributes of challenge that players could articulate in interviews and which would, therefore, qualify as a basis for items in a questionnaire to measure the experience of perceived challenge. As we had a clear focus around the elucidation of the felt experience of challenge, an iteratively developed *open coding* (Blandford, 2013), more akin to *thematic analysis* (Braun and Clarke, 2006), was used to synthesise a description of the data rather than taking a more theoretical account as might be offered by e.g. grounded theory (Charmaz, 2014; Salisbury and Cole, 2016).

3.1. Interview Methodology

We conducted semi-structured interviews with four players and five games user researchers ($N = 9$, 4 women and 5 men) with varying levels of gaming experience, game preferences and skills. Based on our aims, sample specificity, use of established theory, quality of dialogue, and analysis strategy (Malterud et al., 2016) this sample size was deemed suitable.

Our participants reported to play both single- and multiplayer games, including role-playing games (RPGs) (e.g. *The Elder Scrolls V: Skyrim*, *Dragon Age: Inquisition*, *The Legend of Zelda: A Link Between Worlds*), massive multiplayer online RPGs (MMORPGs), massive online battle arena games (MOBAs), shooters (e.g. *Counterstrike* as well as *Overwatch*), sport, racing, adventure (e.g. *Firewatch*, *That Dragon, Cancer*), action adventure (e.g. *Shadow of the Colossus*, *Prince of Persia*), puzzle (e.g. *Bejeweled*, *2048*), strategy (e.g. *Civilization*, *XCOM*), simulation (e.g. *Prison Architect*, *Stardew Valley*), musical (*Audiosurf*, *Guitar Hero*), arcade (e.g. *DoDonPachi Daifukkatsu*) and a range of other games. This diversity allowed us to obtain relevant information about their perceptions of challenge in different game types and to relate to different skill-sets.

Each interview began with participants describing their favourite games and commenting on whether they considered them challenging. The follow-up questions aimed at collecting data related to the players' perception of different types of challenge. We did not provide any definitions or examples of challenge types to avoid confirmation bias. Participants were asked to elaborate on their opinions regarding different types of challenge in the games they play and describe how these experiences made them feel. On average, the interviews took 33 minutes.

3.2. Interview Analysis

Open coding (Blandford, 2013) was deployed as a method for identifying and reporting components of challenge present within our participants' interview data. Our initial coding system was developed prior and independently to the publication of the work by Bopp et al. (2018). We consider any overlap with their codes and similarities in the identified themes as support for our findings. We first transcribed the interview data, studied it in depth and then iteratively assigned codes and clustered them into themes that provided overarching relationships between the individual codes. For this, we used the traditional challenge categories, as described by Adams (2014), which were complemented with the additional category of emotional

challenge proposed by Cole et al. (2015). To ensure consistency, the database of codes was shared between the two coders: the first author and an independent researcher. The final coding system consisted of 4 main categories that we call here themes. These are *cognitive challenge*, *performative challenge*, *emotional challenge*, and an additional theme that was related to all three types: *management of challenge in games*. The latter primarily focused on the relationships between challenge and other experiences.

Players reported their experiences of challenge in respect to a broad variety of video games, some of which we mentioned earlier. When talking about their expectations towards challenge in games, participants had mixed responses – some were actively looking to test their skills against the challenges that games offer, while others preferred more relaxing sessions. This reflects the diverse preferences described in the earlier literature.

Most players reported challenge to be an important part of their gaming experience: *"It's nice to be challenged. Because if it was too easy, then it would be kind of boring, I suppose"* (P1). They also remarked upon challenge being a distinguishing feature of video games as opposed to films. The experience of challenge they described as being desirable was aligned with the notion of 'balance' in *flow* (Csikszentmihalyi, 1990): *"stuff that raises your interest levels enough, but it doesn't put you off from continuing"* (P9).

Consistent with the previously discussed research, the players referred to the three main types of challenge: performative, cognitive and emotional. We report the results separately, with reference to the corresponding games, where appropriate.

3.2.1. Cognitive Challenge

The first few occasions of playing a new game present particular cognitive challenges, some of which can only be solved through trial and error: *"knowing what weapons to pick up, and the best armour and things, was the initial challenge for me"* (P7 – *PUBG*). Participants also mentioned that their expectations were not always met when choosing a strategy that seems appropriate at that specific time and place: *"Doing stuff that is counter-intuitive to end up with stuff that is correct in the end"* (P4 – *Kami*). At this stage, useful feedback from the game is crucial to understanding the outcome of actions and learning the rules.

Once players have learnt the game basics, planning ahead and trying to anticipate the outcome of their actions becomes important. Being able to focus on the long game is one of the key challenges that many RPGs

and strategy games offer to their players: *"trying to balance your options and making sure I've looked to see in the future what people will be able to do [...] You kind of decide what you want them to be before they actually become that"* (P5 – *Final Fantasy*).

Participants also emphasised when they had to retain certain information in their memory to progress in the game. *"You have to remember who told you that and where to find it roughly, and then physically search for it. [...] If you haven't done it and then you come back, you don't remember what you supposed to do"* (P1 – *Skyrim*).

Similarly, participants also described situations in which the game required them to memorise and keep track of multiple dimensions, e.g. managing game resources: *"you have to grow plants, you have to satisfy your customers, and also remembering what customers like"* (P4 – *Weed Firm*). Players deal with this "invisible complexity" by keeping track of multiple tasks at the same time, looking for subtle cues, and constantly checking the state of the game to prevent it from spinning out of control. This complexity is exacerbated by limited resources: *"you either have to grind a lot and there are not many grinding [performing repetitive tasks] spots in the game, or you have to really conserve your resources because those are also limited in the game world"* (P8 – *Lisa: The Painful*). This includes the resources needed to concentrate and recall the game from one session to the next: *"more likely to just start off fresh, build back up again, rather than continue"* (P1 – *Civilization*). Team members (real or virtual) can be another resource that need to be managed and planned for: *"deciding how you are going to level each character up, what skills you are going to give them, and making sure you've got a balanced party"* (P5 – *Final Fantasy*).

Challenge can come simply from requiring persistence or "brute force". However, games which in contrast require a strategic approach that requires planning, evaluation and decision-making were perceived as more challenging: *"Any situation that you can't brute force yourself through, that involves figuring out a strategy that works better or best on managing resources"* (P8).

3.2.2. Performative Challenge

Performative challenge, however, is not the same as "brute force" and was not viewed as the easier option. Physically challenging games test a different set of players' skills. Participants reported dexterity, quick reflexes, reaction time, spatial awareness and reasoning, mechanical or muscle memory, stamina, persistence, and hand-eye coordination as the skills associated

with performative challenge.

In contrast to the previous focus on mentally demanding long-term planning, participants stressed the need to make snap decisions in real-time: *"the longer you play the more difficult it gets. So, you've got to move things faster and you've got to turn things and fit things into other things, and that becomes more physically and mentally challenging, as the game progresses"* (P5 – *Tetris*).

Dexterity was another frequently mentioned skill when talking about physical challenge. It requires players to react not only fast but also accurately: *"You have to react really quickly to what's going on, and you've got to be able to physically press the buttons fast enough and move things in the right way"* (P5 – *Tetris*). P8 reported that it was *"hard to pull off certain moves"*, which added to their experienced challenge. Some games also require players to learn and recite 'combo' moves, which tests not only players' physical skills, like dexterity, but also their cognitive skills, such as memory.

3.2.3. Emotional Challenge

The final challenge described in the interviews is emotional challenge. Unlike cognitive and physical challenge, emotional challenge arises as a result of dealing with difficult content based on moral choices and their consequences, relatedness, empathy and attachment to the characters in the game. Our participants reported on emotional experiences from following the story and getting immersed in the narrative of the game: *"Trying to find that one thing that joins all the dots together. So that's challenging in a good way"* (P9 – *Firewatch*).

One of the main themes reported in the context of emotional challenge is moral dilemmas (or difficult choices) and their consequences. These choices are perceived as difficult for several reasons. In some cases, the consequences of making a choice are unknown and, therefore, the uncertainty of making such a decision is challenging. Typically, the given options are purposefully difficult to choose amongst, as none of them appear to be morally desirable: *"the choice is you kill the demon which would solve the problem, but by doing so you would kill the child. And another option is you could leave, which of course wouldn't help you. Or the other one is [. . .] to kill the demon without killing the child, but to do that you have to kill . . . the child's mother"* (P1 – *Dragon Age: Origins*).

Role-playing was reported as a major part of playing emotionally chal-

lenging games. Players adopt a character and then play out that character. Deciding how the character should act forms part of the emotional challenge: *"If you play as what you consider a morally good character, and you consider yourself to be a morally good person, [the personalities and decisions] become intermingled quite quickly. Whereas if you decide to play as a 'tank', and then he just kills everybody, and that's what you want to do because it's fun, then it's not such a big deal. But then it's not particularly emotionally challenging"* (P5 – *Skyrim*). Some dilemmas are not so much about character but bringing to the fore whether the game is a game or a story: *"You get caught by another gang and then you get the choice: either they kill all your party members at the time that are with you or they cut off your arm. [. . .] If you choose to lose your arm, that changes all your combos that you can do, or, obviously, you lose all your party members, which is also really shitty, but you have to choose between the two of them"* (P8 – *Lisa: The Painful*) However, having role-played once, players can step back from the emotional challenge to explore the game further: *"The first time I went through [the game], I was playing in a more emotionally invested way, but if I were to play through it again, I would probably be playing it in a more curious as to what the other options result in sort of way"* (P7 – *Life is Strange*).

Life is Strange emerged as a game that provoked emotional challenge by allowing players to assess themselves against the decisions and actions of other players: *"There is this sort of judgement in my mind, did I make a [. . .] a controversial choice, or morally 'Am I weird?' kind of choice?"* (P7). Similarly to physically or cognitively challenging games, participants enjoyed discovering something new about themselves. However, here, they appreciated learning more about their views and perceptions rather than physical or mental capabilities: *"One thing about challenge is that it helps you measure [your] identity and create that identity"* (P6).

Moral dilemmas were viewed positively, with participants stressing that *"you get invested into it"* (P1). However, choosing between 'less than ideal' options led to the challenging feeling of guilt: *"I remember that was really hard, I was sitting there for at least 5 minutes, which doesn't seem like a lot, but in the game world that's like eternity. I was like: 'What do I do now?', and then I chose to rescue the party members and the girl lost her finger. And even to this day, I don't feel good about this decision."* (P8 – *Lisa: The Painful*). Players' attachment to the game characters becomes particularly difficult when playing games with the 'permanent death' mechanic: *"when your character dies – they die, and you feel it"* (P6 – *XCOM*).

Guilt, however, was not the only emotional response participants reported when talking about their interactions with other characters in the game. They also positively referred to their attachment to the virtual characters: *"You got these romance plots and you feel for the bond that the characters have for each other. It is very sweet."* (P1 – *Dragon Age: Inquisition*). Bonding with characters made participants feel a *"sense of obligation towards party members"* (P8 – *Lisa: The Painful*), which made in-game choices more meaningful and consequently harder. Such hard decisions were valued as making the game relevant and "feel more real": *"it was less abstracted from real life, so you could imagine how that must feel for the characters. [...] Finding the consequences to something like that – it's quite emotional"* (P7 – *Life is Strange*).

3.3. Managing Challenge

In addition to the three distinct challenge types, we also identified several constructs that were related to challenge, but cannot be classified as part of this experience. Instead, these can be viewed as experiences that influence or are influenced by perceived challenge.

The assumption that players learn during play is at the heart of all three challenge types, as they require players to evolve, hone and master their skills in order to make progress. But learning also mediates the *perception* of challenge: participants associated learning and a sense of achievement or accomplishment, which in itself was positive: *"I know it's hard and I feel really good when I finish it"* (P2 – *Dark Souls*). Participants also enjoyed learning about new strategies and were curious about their own capabilities: *"I was interested whether I could do it and I was pleasantly surprised"* (P6). This *"sense of self-improvement"* was one of the key motivators that participants named when talking about why they choose to play challenging games. According to Malone (1981), challenge could be viewed as curiosity about the player's abilities, or curiosity can be explained as a challenge to one's understanding. Similarly, a sense of self-improvement closely relates to changes in "competence" as component of the SDT of (player) motivation (cf. Ryan et al. (2006)).

Useful feedback is paramount to learning (Gee, 2008): In order to progress in a game, players need to be able to understand why they succeed or fail. Players feel stuck when being unsure of the cause of their failure, unable to find any issues with their approach, or when feeling like they have exhausted all their strategies: *"There is a certain strategy you do for the best time limits*

and move limits, I just couldn't do it no matter how many times I did it." (P2 – *Candy Crush*).

Many games require players to practice their skills in order to overcome challenges and progress through the game. This has been reported with respect to cognitive skills / challenges: *"I prepared for this maybe a week, in the evening, always grinding and strategising and trying different stuff to fight those bosses"* (P8 – *Final Fantasy*). However, it also holds for physical skills / challenges: *"as you get better, you keep going forward, but if you practiced a lot, you become really good at the game. You don't need to have more knowledge about the game"* (P2 – *Guitar Hero*). This emphasises that time and effort are a crucial requirement for learning in games.

Repetition of the same challenge is often the only way to practice: *"If I wanted to get to the bottom of the mine, I had to build up my character, and kind of level it up, so I had to keeping coming back to the mine. That was quite challenging"* (P7 – *Stardew Valley*). Grinding though can be repetition without skill or game progression: *"you just have to go back and dungeon crawl and get lots of XP"* (P5 – *Final Fantasy*). Such persistence tended to be driven by players' curiosity to see how the game progresses but without novel aspects, repetition can become boring or frustrating. This eventually can lead to disengagement, i.e. player churn: *"the longer you play the better you get at the time, and you feel the sense of achievement, but you haven't achieved anything – you have just been playing the game long enough to get that new item to make you better."* (P2 – *Stardew Valley*). Novelty or interest can be maintained by varying the task at hand: *"I don't like being stuck in one area and just not being able to get past it. It just makes me feel incompetent in a way ... Whereas if you got different routes of it, or different options for solving different puzzles and things, that would make challenge more fun"* (P7).

Frustration can also arise from a loss of progress, such as having to restart a level, or when progress is prevented by and issue outside the player's control. For example, waiting for an action to be completed is common in many mobile games as part of the micro transactions mechanics (e.g. *SimCity BuiltIt*). Similarly, games like *Candy Crush* allow players to buy hints in order to progress in the game. Participants described this kind of mechanics as unrelated to their skills and this, hence, is *not* perceived as exciting or challenging: *"A lot of the times I get frustrated, I feel like games game me. [...] You're making this level particularly hard because you're trying to manipulate me into buying this extra stuff"* (P4 – *Bejeweled*).

Another reason for losing interest in a game is insufficient challenge in the first place, or because it has been overcome: *"you don't want to play it anymore because you know how it works"* (P4 – *Plague Games*). This was especially true of games that allow the player to master them – this happens when players' skills cannot be improved further, and the game does not offer any new challenges. One way to keep the challenge fresh is to play with other humans and train substitute skills: *"if all the other players were bots, for example, I don't think it would be as challenging as I think the human element is [...] There are so many different strategies that you could be on with"* (P7 – *PUBG*).

3.4. Discussion of the Findings

Our findings from the interviews complement Bopp et al. (2018)'s work in many respects. There, players also referred to cognitive and physical challenges very explicitly. Furthermore, their findings also cover our interviewees' experiences of difficult decisions (Decisions & Actions; Difficult Themes) and facing both the game and emotional consequences of those decisions (Intense Emotion). In contrast to Bopp et al. (2018), however, playing socially was not reported as a distinct type of challenge in itself by our participants, but rather as making the other types of challenge more meaningful. According to their research, other players contribute to challenge in that they add meaning and a potential emotional outcome over and above that presented by the game itself. This is not to say that players are not challenged by this, but rather that the challenge experienced is not intrinsic to the game but relational, and perhaps works to amplify the sensation of challenge that is encoded in the game.

Our interview study thus shows commonality with both the literature and the previous empirical study of challenge in games. In addition, the varying ways in which challenge is made manifest to our participants served as an important, additional source for shaping our questionnaire.

4. Stage I: Item Generation and Scale Construction

We created an initial items pool based on the survey of relevant literature and interviews with players in Section 2 and 3, respectively. In particular, for each of the types of challenge identified in the interviews, we crafted items based on what players said about that type and tried to cover the major aspects that players referred to. This was complemented by additional items

based on the accounts of challenge from the semi-structured qualitative study. At this early stage, we did not avoid words or phrasings that might be close in meaning or overlap in wording, in order to achieve maximally broad coverage.

We deliberately did not include items related to social connections between players, for three reasons. Firstly, social connection involves a rich range of meanings, only some of which might be related to challenge. Secondly, connection to other players specifically has already been studied separately and there exist reliable questionnaires that address these connections in detail (Hudson and Cairns, 2014; De Kort et al., 2007). Thirdly, in our interviews, social play seemed to make the experienced challenge more meaningful, but did not challenge the players more directly over and above the identified experiences of cognitive, physical and emotional challenge. We thus refrained from trying to measure social challenge as separate construct.

These steps resulted in 120 items. We refined this preliminary item pool by eliminating items that were not deemed relevant to one's experience of challenge, or appeared too similar in their meaning or phrasing. We also removed any items related to other gaming experiences, such as uncertainty (Power et al., 2018), immersion (Jennett et al., 2008), or competence (Ryan et al., 2006), as these are covered in dedicated instruments.

This resulted in a refined, yet early pool of 60 items (Table 2), covering different aspects of challenge. They address the three types of emotional, cognitive, and physical challenge, as well as further themes such as perceived evaluation of performance in the game, perceived difficulty of the game, and experiences related to managing challenge in games.

5. About Our Approach

The remainder of this paper is about providing statistical support for the concept of challenge reified through this set of items. This is commonly done through factor analysis (Kline, 2014) and in packages like SPSS and SAS, factor analysis methods are primarily based on what is known as Classical Test Theory (CTT). This approach is essentially what is described in textbooks like Nunally and Bernstein (1978), but it makes certain statistical assumptions including glossing over the fact that Likert items are not continuous measures and that different items differ in both their sensitivity and accuracy as measures of the underlying concept. These used to be necessary simplifying assumptions to make factor analysis computationally tractable.

However, with increasing computing power, more sophisticated approaches are now possible.

In this paper, we, therefore, use Item Response Theory (IRT). This approach is beginning to gain ground in psychology (Embretson and Reise, 2000) and more recently in human-computer interaction (Cairns, 2019, Chapter 16). IRT overcomes the limitations of CTT by treating Likert items as discrete, ordinal measures that differ in the sensitivity and accuracy of the underlying latent concepts. In particular, because it does not make simplifying assumptions about the relationship between items and the latent concepts, IRT is not adversely affected by distributional problems, such as skew or varying means and standard deviations. Fortunately, for practical purposes, the results of using IRT to do factor analysis are broadly the same, namely, a set of factors reified in the instrument and a set of loadings that show how each item relates to the different factors. In the analysis presented here, we used the *mirt* R package (Chalmers et al., 2012) which is able to perform multidimensional IRT. The rotation method used to achieve simple structure is Direct Oblimin.

We have not used a structural equation modelling (SEM) approach because this would presuppose a specific hypothesis about the structure of the different concepts underlying challenge. Instead, we are using an exploratory approach to determine a plausible factor model that is then confirmed in a further study. With IRT, there are not the obvious measures of model fit, such as found in CTT-based methods, and indeed such measures are more suited to the comparison of hypothesised models than definitive indicators of the quality of a model. Instead, the primary statistics used to evaluate the proposed factor structures are first α and item-drop correlations. Cronbach's α is the usual measure of internal consistency that indicates the degree to which items of a factor relate to each other and therefore to a putative latent concept. While there are alternative measures of internal consistency that have fewer assumptions (Dunn et al., 2014), we found that they did not give meaningfully different interpretations and, therefore, we report Cronbach's α as the measure of internal consistency of the subscales.

Item-scale correlations are the correlations between an item and the corresponding factor score (taken as a sum of the constituent item scores) to which the item belongs. However, because items are part of that score, this gives an immediate correlation between items and factors scores. Thus, a better measure of the relationship between an item and its factor is the item-drop correlation where the item is correlated with the factor dropping the

item itself from the score. A relatively high item-drop correlation indicates that an item is measuring the same concept as the remaining items of the factor. Ideally, in IRT, we would use the test information function (Thomas, 2011) that indicates the amount of information each item contributes to the factor. However, in this particular context, the interpretation of the information test functions coincided with that of the item-drop correlations, so these are instead used as being a more familiar concept.

6. Stage II: Exploring the Underlying Factor Structure of the COR-GIS

We conducted a survey study in order to perform an exploratory factor analysis (EFA) on the initial pool of 60 items. The analysis was then used to shorten the initial set of items so they could be used as a more practical yet valid instrument.

6.1. Data Collection

The 60-item questionnaire was administered using an online survey. In order to gather responses from a diverse audience of digital game players, we distributed the survey through various social media channels and online gaming forums, including *Twitter* and *Reddit*. Each participant was briefed on the usage of their data in accordance with the ethical clearance provided on the study.

We asked our participants to reflect on their most recent experience of playing a digital game and to choose answers that best reflected their experience. All items had a 7-point Likert scale anchored at the ends with 'Strongly Disagree' and 'Strongly Agree'. The position of the questions was randomised for each participant to avoid order-effects.

Overall, 432 respondents participated in the survey. Not all participants answered every item: 38 participants had left 28 or more ($> 45\%$) challenge items blank, and hence were removed from the analysis. 42 further participants had not answered between 6 and 22 (10% and 37%) of the items, but these were left in as we could not discover a pattern of non-response. Of the $N = 394$ included respondents, there were 57 women, 332 men, and 5 non-binary or not specified. The average age of the participants was 25.9 years (SD = 7.2) and the average number of gaming years was 18.2 (SD = 6.8).

Genre	No of Games
Action RPG	137
Action-adventure	47
Adventure	12
Beat 'em up	2
Eroge	1
Fighting	2
Grand Strategy	4
Management Simulation	5
MMORPG	13
MOBA	11
Platformer	11
Puzzle	2
Roguelike	1
Role-playing	68
Shooter	61
Sports	2
Stealth	1
Survival	1
Survival horror	1
Tactical RPG	7
Turn-Based Strategy	1
Vehicle simulation	4

Table 1: Stage II: Frequency of video game genres played by the survey respondents in the first study.

Respondents reported playing over 100 titles, with some of the most popular being *Nier Automata* (43), *Persona 5* (38), *Horizon Zero Dawn* (20), *Overwatch* (17), and *The Witcher 3* (12) (cf. supplementary materials for a full list of games: https://doi.org/10.25383/city.9221843). Overall, they played a wide range of genres, which are listed in Table 1.

6.2. Item Analysis

Individual items had typical response distributions for experience questionnaires of this sort: the means were typically between 4.5 and 6.5 and thus on the agreement side, with SDs between 1 and 2.25. There was one exceptional item which was strongly negatively skewed and had a SD of 0.9. Medians were mostly either 5 or 6 showing a degree of negative skew. No medians were at the extreme ends of the scale, showing that there was no instances of very strong skew. Only the item `#54` "My actions in the game made me guilty" had a median of 2 ('disagree') which, given the strong wording of this item, is not too surprising. No item exhibited bimodality. Because IRT is robust to variations in distributions compared to CTT, no items were removed based on their response distributions.

Following standard analysis procedures (Hair et al., 1998), the Kaiser-Meyer-Olkin (KMO) Measures of Sampling Adequacy (MSA) was considered for both the whole dataset and individual items. The full data had MSA = 0.92 and no item had an individual MSA < 0.7. This suggests that no item was unsuitable for factor analysis.

6.3. Exploratory Factor Analysis

A scree plot with parallel analysis (Horn, 1965; Cairns, 2019) suggested there were around 5 factors in the data. However, as challenge is often referred to as a single concept, a single factor solution was generated and is presented in Table 2. Using a loading cut-off of 0.35, which is a common threshold for assigning items to factors (Hair et al., 1998), 43 of the 60 items loaded on this single factor. This does suggest that while the sources of challenge may be diverse, they do, to a degree, relate to a single underlying concept of challenge. However, as only two items loaded above 0.7, this suggests that the coherence of the experience as a single factor across all types of challenge is only modest.

Accordingly, 4, 5 and 6 factor solutions were explored. Initial components were rotated using Direct Oblimin to achieve a simple factor structure. The 6 factor solution was similar to the 5 factor solution, but with one extra factor of only a few items that strongly loaded on it and that also cross-loaded with other factors. Both the 4 and 5 factor solutions had reasonable simple structure and there was conceptual coherence to the items in each factor. The main difference between the two solutions was that a factor in the 4 factor model that could be reasonably named 'cognitive challenge' was split across two factors in the 5 factor solution with the additional factor corresponding to the planning component of cognitive challenge as highlighted in the previous sections. Understandably, however, these two factors in the 5 factor solution correlated ($r = -0.38$). We felt that planning and cognitive challenge were sufficiently close that the 4 factor solution still gave a strong account of the data, while reducing correlation between factors.

#	Item	Single	COG	EMO	PERF	DM
1	Succeeding in the game required much planning	0.59	-0.80			
2	I had to memorise a lot of different things when playing the game	0.52	-0.71			
3	Managing time in the game required much planning	0.45	-0.69			
4	Managing resources in the game required much planning	0.35	-0.67			
5	I had to remember and recall much information in the game to be able to play well	0.58	-0.66			
6	I had to think several steps ahead when playing the game	0.63	-0.65			
7	I had to prepare for the things that the game threw at me	0.62	-0.60			
8	The game made me manage several tasks at the same time	0.58	-0.57			
9	I had lots of different things to think about at once in the game	0.66	-0.56			
10	Playing the game requires great effort	0.71	-0.54			
11	I had to measure the benefits of each thing I did in the game	0.60	-0.53			0.39
12	I felt challenged when playing the game	0.69	-0.51			
13	I had to constantly keep track of what was going on in the game	0.65	-0.46			
14	Playing the game was demanding	0.65	-0.46			
15	I was always refining my tactics in the game	0.55	-0.45			
16	Playing the game felt like solving a puzzle	0.34	-0.45			
17	I had to think actively when playing the game	0.72	-0.43		0.36	
18	Playing the game required me to do my best	0.69	-0.43		0.42	
19	Playing the game requires determination	0.69	-0.42			
20	I couldn't brute force my way through the game	0.40	-0.40			
21	Time pressure was an important part of playing the game	0.48	-0.39			
22	Grinding was an important part of the game	0.09	-0.39			
23	I had to do things within set periods of time when playing the game	0.42	-0.35			
24	Caring about characters was an important part of my experience of the game	0.06		-0.84		
25	This game is more than just a game to me	0.39		-0.76		
26	The things that happened in the game made me sad	0.29		-0.76		
27	I felt a sense of responsibility for characters and events in the game	0.33		-0.69		
28	The game made me think about real life issues	0.21		-0.69		
29	I invested much thought into the game	0.52		-0.60		
30	Dealing with negative emotions is an important element of this game	0.33		-0.56		
31	Playing the game was stimulating	0.46		-0.48		
32	I think about my in game decisions even when I am not playing	0.43		-0.47		
33	I felt a sense of suspense when playing the game	0.55		-0.47		
34	The game had moral dilemmas in it where the choice was not obvious	0.17		-0.46		0.43
35	The game involved making moral choices that I didn't agree with	0.11		-0.37		0.41
36	I had to react quickly when playing the game	0.63			0.97	
37	I had to act quickly when playing the game	0.63			0.95	
38	Thinking fast was an important part of the game	0.63			0.91	
39	Quickly responding to things that I saw was an important part of the game	0.60			0.84	
40	I had to make snap decisions when playing the game	0.63			0.76	
41	Playing the game required timing my actions	0.58			0.74	
42	Playing the game demanded precision in my actions	0.65			0.71	
43	Being accurate was important in the game	0.62			0.69	
44	I had to be accurate in my actions when playing the game	0.66			0.59	
45	The game kept me on my toes	0.64			0.57	
46	There were some decisions in the game that I regretted	0.34				0.77
47	I wonder how different the outcome in the game would be had I chosen a different option	0.32				0.72
48	I had to make difficult choices in the game	0.49				0.70
49	Progressing in the game involved making difficult choices	0.51				0.67
50	When faced with decisions in the game the choice was obvious	-0.19				-0.58
51	The game made me think hard about my decisions	0.55				0.56
52	I had to think about possible alternatives for my actions in the game	0.50				0.55
53	I had to predict the effects of the things I did in the game	0.53				0.51
54	My actions in the game made me anxious	0.45				0.39
55	I spent much effort predicting what was going to happen next in the game	0.53				
56	The game didn't give me much warm up time	0.09				
57	I often couldn't decide what to do next in the game	0.15				
58	My actions in the game made me guilty	0.24				
59	The problems presented to me in the game were easy to solve	-0.41				
60	The game was more about doing things than figuring them out	-0.18				

Table 2: Initial items pool and the four derived factors from IRT analysis in the first study. Only loadings > 0.35 are shown.

Seven items were excluded post-analysis due to their low factor loadings on any of the 4 factors (items `#54-60` (below the dashed line) in Table 2). The remaining items factored well into 4 components which we defined as:

- **Cognitive Challenge** (COG): arising from the need for preparation, planning ahead, memorisation, effort and multi-tasking.
- **Performative Challenge** (PERF): arising from the game requiring rapid and accurate action from the player.
- **Emotional Challenge** (EMO): arising from the emotions evoked in the player which might also have implications for things they thought about outside of the game.
- **Decision-Making Challenge** (DM): arising from having to make choices that were difficult or could lead to regrettable outcomes.

To refine the questionnaire down to a more practical size, items were removed primarily based on loading so those items with the highest loadings were retained. Within the IRT framework, it is appropriate to look at the contribution of information of each item to its factor. Analysis of item information agreed with the decisions based on loadings.

In some cases where loadings were similar, we kept the items that were more applicable to a wide range of games. Where the wording of items was similar, a decision was made whether the difference in wording was sufficient to keep the items, taking into account also the strength of the loading.

Two items (`#34`,`#35`) were retained in the EMO factor, although they had relatively low loadings and cross-loaded with the DM factor. This is because we felt that they represented a very specific type of challenge, namely moral dilemmas, which showed in our interview analysis but was not captured elsewhere in the questionnaire yet. For this reason, they were included in the final set of EMO items. However, we recognise that in some circumstances, these items ought to be omitted as being inappropriate to some games.

The refinement process resulted in the final pool of 30 items (Table 5 lists the reduced item set) measuring **Cognitive**, **Emotional**, **Performative**, and **Decision-Making Challenge**. Cronbach's α for each reduced factor on this dataset is 0.88 (11 items), 0.86 (9 items), 0.93 (5 items) and 0.84 (5 items) respectively, indicating very good internal consistency.

The item-drop correlations for each item in the COG scale were between 0.5 and 0.65, for PERF scale 0.7 and 0.87, for the DM scale, 0.55 and 0.7. For the EMO scale, all items except one had item-drop correlations between 0.5 and 0.75 except for one item which had a correlation of 0.39. On the whole, these, therefore, show that each item was relevant to the remainder of its scale whilst also not so highly correlated as to indicate redundancy. The exceptional item in the EMO scale was "Playing the game was stimulating" however this correlation is not so low as to be irrelevant and does reflect an arousal component of emotion as well as loading reasonably well in the initial analysis. We, therefore, felt all items reflected a coherent contribution to their corresponding subscales of the reduced 30-item questionnaire.

r	**COG**	**EMO**	**PERF**	**DM**	α
COG	1	0.25	0.44	0.38	0.88
EMO		1	0.06	0.56	0.86
PERF			1	0.01	0.93
DM				1	0.84

Table 3: Pearson's correlation r between the subscale scores and internal consistency (Cronbach's α) of the reduced 30-item set.

The reduced item scales were scored by averaging the scores of the constituent items. The correlations between the factor scores as well as consistency ratings per factor are given in Table 3. The consistency ratings suggest good internal consistency to each factor. The degree of correlation between the factors is somewhat high suggesting that there is some conceptual overlap in the experience of challenge. This is consistent with the relatively coherent single factor solution found initially. However, these items were also chosen for reduced cross-loading suggesting that though there is a commonality in the experience of challenge there is a nuance in the types of challenge that is worth capturing.

Therefore, as an initial set of factors, this seems to be a plausible model of challenge with four distinct components that are inter-related but also show some degree of discrimination between the different types of challenge.

7. Stage III: Confirming the Factor Structure of the CORGIS

We conducted a second online survey to perform a confirmatory factor analysis (CFA) on the final 30-item questionnaire. Our goal was to verify whether the trimmed-down item set produced the same factors thereby

indicating the robustness of the model as capturing persistent concepts of challenge in digital games.

7.1. Data Collection

The survey was distributed on several online channels, including *Twitter* and *Reddit*, to gather responses from the most diverse gaming audience possible. Each participant was briefed about the ethical implications of the study, after which they were asked to complete a demographics questionnaire. Respondents were then asked to pick a game they recently played and answer the survey based on their experience of that particular, recent game-playing session. In the online survey, all items were ordered randomly for each participant.

This time, 1001 video game players responded to the survey. Thirteen respondents were removed as they had not given answers to five or more ($> 15\%$) of items. Only 48 more respondents had failed to give complete responses. One respondent had assigned the same maximum value to every single item and was removed as they likely did not deliberate their responses. There was no pattern of non-response by item. Of the $N = 987$ included respondents, there were 83 women, 874 men, and 30 non-binary or not specified. Their average age was 24.62 (SD = 6.48) years and the participants on average have spent 17.49 hours per week playing video games (SD = 6.70).

Amongst the most popular games played recently by the participants were *Monster Hunter World* (89), the *XCOM* series (130), the *Civilization* series (66), *Life is Strange* and *Life is Strange: Before the Storm* (73), the *Shin Megami Tensei* series including *Persona 5* (114), and the *Mass Effect* series (54) (cf. supplementary materials for a full list: https://doi.org/10.25383/city.9221843). Overall, titles came from diverse genres, which are listed in Table 4.

7.2. Confirmatory Factor Analysis

To further evaluate construct validity of the new questionnaire, we split the participants' data as follows. Players of the three games *Monster Hunter World*, *XCOM* and *Life is Strange* were randomly divided into two halves. These games were chosen because they all had a large number of players but very different play styles which we would expect to appear as differences in the experience of challenge. Half of the players from each game was held out of the analysis and used to compare and provide validation for the challenge experience as described in the next section. The other half were combined

Genre	No of Games
Action RPG	258
Action-adventure	39
Adventure	89
Arcade	1
Beat 'em up	5
Collectible card game	5
Fighting	12
Grand Strategy	25
Life Simulation	3
Management Simulation	7
MMORPG	10
MOBA	14
Platformer	13
Puzzle	4
Real-Time Strategy	5
Rhythm	3
Roguelike	10
Role-playing	159
Shooter	79
Sports	2
Survival horror	5
Tactical RPG	98
Turn-Based Strategy	134
Vehicle simulation	7

Table 4: Stage III: Frequency of video game genres played by the survey respondents in the second study.

with the remaining data to make up a total of 839 participants in the confirmatory factor analysis.

As before, the KMO MSA was checked. The MSA for the full dataset was 0.93 and no item had an MSA < 0.85 which indicates that the full set of items was suitable for factor analysis. A scree plot with parallel analysis (Horn, 1965) suggested 4 factors with the "knee" in the scree plot appearing on the fourth factor with an eigenvalue of 1.21 just below the parallel analysis value of 1.25. This matches the expectation of four factors based on the model.

Since the previous data and analysis had suggested a possible single dimension of challenge across the entire set of items, a single factor solution was generated with loadings shown in Table 5. Unlike in the previous study though, there was a less clear picture of a single coherent notion of challenge. The DM items all loaded reasonably well (above 0.59) on the single factor, but no PERF items loaded on it (below 0.25). The EMO items all loaded above 0.48 but only 2 above 0.6, and the COG items loaded weakly with 6 items loading below 0.5 and only 2 above 0.6. This suggests that the reduced questionnaire does not represent a single coherent concept of challenge. In-

Subscale	Item	Single	COG	EMO	PERF	DM
Cognitive Challenge	1. Succeeding in the game required much planning	0.56	0.76			
	2. I had to memorise a lot of different things when playing the game	0.38	0.57			
	3. I had to think several steps ahead when playing the game	0.62	0.77			
	4. I had to prepare for the things that the game threw at me	0.58	0.71			
	5. Playing the game requires great effort	0.44	0.71			
	6. I felt challenged when playing the game	0.49	0.77			
	7. I had lots of different things to think about at once in the game	0.60	0.71			
	8. The game made me manage several tasks at the same time	0.43	0.59			
	9. I had to constantly keep track of what was going on in the game	0.45	0.67			
	10. I had to think actively when playing the game	0.57	0.76			
	11. Playing the game required me to do my best	0.47	0.72			
Emotional Challenge	1. This game is more than just a game to me	0.46		0.66		
	2. The things that happened in the game made me sad	0.48		0.79		
	3. I invested much thought into the game	0.77		0.59		
	4. I felt a sense of responsibility for characters and events in the game	0.62		0.81		
	5. The game made me think about real life issues	0.51		0.80		
	6. Playing the game was stimulating	0.54		0.35		
	7. I felt a sense of suspense when playing the game	0.58		0.48		
	8. The game had moral dilemmas in it where the choice was not obvious	0.56		0.85		
	9. The game involved making moral choices that I didn't agree with	0.43		0.69		
Performative Challenge	1. I had to react quickly when playing the game	-0.11			0.97	
	2. I had to act quickly when playing the game	-0.09			0.97	
	3. Thinking fast was an important part of the game	-0.01			0.93	
	4. Quickly responding to things that I saw was an important part of the game	0.01			0.89	
	5. I had to make snap decisions when playing the game	0.10			0.83	
Decision Making Challenge	1. There were some decisions in the game that I regretted	0.59				0.68
	2. I wonder how different the outcome in the game would be had I chosen a different option	0.69				0.77
	3. I had to make difficult choices in the game	0.80				0.87
	4. I had to think about possible alternatives for my actions in the game	0.73				0.76
	5. The game made me think hard about my decisions	0.84				0.88

Table 5: Final version of CORGIS with reduced item pool: single factor and confirmatory four factor loadings from the second study.

stead, there are distinct types of challenge being experienced by the players, as suggested by the literature and our exploratory study.

A CFA was done on the data using the 4-factor structure found in the previous study, as the model to fit with items not related to a factor coerced to have 0 loading. This indicates good support for the proposed model of challenge factors. The loadings are also shown in Table 5. PERF has high loadings (above 0.8) of every item. The DM and COG items also have good loadings, though one COG item's loading dips below 0.6. The EMO items also load reasonably well but for the item "Playing the game was stimulating" that loads below 0.35 and the item "I felt a sense of suspense when playing the game" that loads relatively weakly at only 0.48. The 4-factor model accounts for 81.6% of the variance in the data.

Item-drop correlations also show a very similar pattern to the previous study: for each item in the COG scale these were between 0.5 and 0.67, for PERF scale 0.76 and 0.89, for the DM scale, 0.60 and 0.75. For the EMO scale, all items except one had item-drop correlations between 0.45 and 0.71 except for the item "Playing the game was stimulating" which had a correlation of 0.32. It may be that this item, which also previously had a low correlation, is not strongly related to EMO. However, it does still seem to

contribute to the scale, even if relatively weakly and the internal consistency as measured by α is unchanged if this item is omitted from the factor.

The correlations between the scored scales and their internal consistency values are shown in Table 6. These values are similar to the ones in the previous study. The Cronbach's α is high > 0.8 throughout, suggesting good internal consistency to each factor. Notice that DM still strongly correlates with EMO. This might have been due to the cross-loading moral dilemma questions, however, omissions of these two items from the EMO factor does not substantially alter any of the correlations and the correlation with DM remains almost the same at $r = 0.63$.

r	**COG**	**EMO**	**PERF**	**DM**	α
COG	1	0.20	0.24	0.39	0.89
EMO		1	-0.07	0.65	0.84
PERF			1	-0.21	0.94
DM				1	0.86

Table 6: Pearson's correlation r between the subscale scores and internal consistency (Cronbach's α) of the final CORGIS scale.

This might suggest a lack of discriminant validity with the two factors conceptually overlapping and, therefore, not being sufficiently distinct components of challenge. To further investigate the discriminant validity, two steps were taken. First, a confirmatory bifactor analysis was done to explore if there was a general factor of challenge underlying the 4-factor model (Reise, 2012). There was no evidence of a strong general challenge factor with only COG and DM loading all items on the general factor and even then with some quite low loadings. Only some of the EMO items loaded on the general factor and none of the PERF items. Thus, even if there is a relationship between the various factors that make up challenge, each factor is making a unique contribution, though collectively they do build up challenge to some extent.

Secondly, further statistics were calculated, following Hair et al. (1998) to consider convergent and discriminant validity. These were Composite Reliability (CR), Average Variance Extracted (AVE) and Maximum Shared Variance (MSV) of each factor and given in Table 7. The composite reliability asserts the internal consistency of each factor as seen previously with the Cronbach's α values, with CR being above 0.8 for each factor. Further, AVE exceeds Fornell and Larcker (1981)'s rule of thumb threshold of 0.5 except

for EMO, which is just below this threshold. This suggests good convergent validity for each factor. In addition, in each case the AVE was above the MSV, which again Fornell and Larcker (1981) recommends as indicating good discriminant validity.

	CR	AVE	MSV
COG	0.92	0.50	0.16
EMO	0.88	0.47	0.42
PERF	0.96	0.81	0.06
DM	0.89	0.63	0.42

Table 7: Composite Reliability (CR), Average Variance Extracted (AVE) and Maximum Shared Variance (MSV) for each factor of the final CORGIS scale.

Overall, our results suggest that each of the four factors captures some distinct component of experienced challenge though there is some degree to which there is an overall sense of challenge to which all four factors contribute. Moreover, the observation that the factor structure found in the previous study is also supported in this distinct dataset, suggesting that there is robustness and generalisability to the factors. Moreover, each factor shows good convergent validity and discriminant validity as part of the factor structure as a whole.

8. Stage IV: Assessing Criterion Validitation

As the confirmatory analysis provided good support for the four components of challenge underpinning the questionnaire, the data on the three games held out from the factor analysis was analysed in an attempt to link the experience of challenge to the gameplay in these games and, therefore, to the experience of players of those games. The three games are:

1. *Life is Strange*: a graphic adventure game where the player is a character in a story who must protect her town from an impending storm.

2. *Monster Hunter World*: an action RPG with the goal to track down and kill monsters in an open world.

3. *XCOM*: a turn-based tactical game where the player must fight against an alien invasion.

The games have quite different styles and so prior to any analysis, we formulated the following hypotheses related to each aspect of challenge:

1. **Cognitive Challenge (COG)**: all three games have a degree of cognitive challenge in planning and keeping track of the game. However, *XCOM* as a turn-based tactical game should be most cognitively challenging.

2. **Emotional Challenge (EMO)**: *Life is Strange* is primarily a character-driven narrative game and should provide substantial emotional challenge. *Monster Hunter World* has little narrative or characterisation and should have low emotional challenge. *XCOM* does have some elements of character and story that might provoke an emotional response in players, but not at the level of *Life is Strange*.

3. **Performative Challenge (PERF)**: *Life is Strange* is primarily dialogue-driven and should thus not require rapid responses and hence have low performative challenge. *XCOM*, being turn-based, should require less rapid responses but may require accuracy in some aspects. *Monster Hunter World* being a real-time combat game should provide high performative challenge.

4. **Decision-Making Challenge (DM)**: Both *Life is Strange* and *XCOM* require players to make precisely defined decisions in order to progress in the game and we hence expect high decision-making challenge. *Monster Hunter World* is more responsive and decisions do not inherently alter the long-term outcomes of the game. Hence, it should provide low Decision-Making challenge.

	Life is Strange (36)	**Monster Hunter World** (47)	**XCOM** (65)	**F(2,145)**	η^2
Cognitive Challenge	3.59 ± 1.29	5.62 ± 0.78	5.97 ± 0.53	96.8	0.57
Emotional Challenge	6.21 ± 0.84	3.43 ± 0.82	4.60 ± 0.76	123.3	0.63
Performative Challenge	3.07 ± 1.56	6.24 ± 0.82	2.67 ± 1.37	116.0	0.62
Decision-Making Challenge	6.28 ± 0.88	3.33 ± 1.42	6.12 ± 0.83	115.1	0.61

Table 8: $M \pm SD$ and results of ANOVA (all $p < 0.001$) on the four components of challenge for the three games ($N = 148$) in the validation analysis.

The four challenge subscales were scored for each game and their means and standard deviations are summarised in Table 8. The components of challenge did change as hypothesised based on the attributes of the different

games, and all with large effects. Cognitive challenge was lowest in *Life is Strange* and highest in *XCOM*, though unexpectedly, *Monster Hunter World* was also high. This might reflect a strategic aspect of the combat in *Monster Hunter World* that presents a cognitive challenge to the players. By contrast and as predicted, emotional and decision-making challenge were highest in *Life is Strange* and lowest in *Monster Hunter World*. *XCOM* provided players with high decision-making challenge which fits with its turn-based gameplay but an intermediate level of emotional challenge. This potentially reflects that some aspects of the game put emphasis on character interactions alongside a narrative that some players may be responding to emotionally. As expected, the rapid combat gameplay of *Monster Hunter World* presented a high performative challenge compared to the other two games. However, the other two games did not have a very low level of performative challenge suggesting that timely actions were still an important component of the gameplay.

It is worth noting that the pattern of emotional and decision-making challenge differed across the three games with XCOM, in particular, showing a lower level of emotional than decision-making challenge. Thus, though these two factors strongly correlate they are able to discriminate between the experiences of different games in ways that are consistent with our hypothesised experience of the games.

This analysis lends weight to the validity of our questionnaire as the challenge components changed in ways that were predicted based on features of the game. However, it is important to be cautious in this sort of analysis. Many other features of the three games are very different but unaccounted for, and it is possible that players might be responding to aspects of aesthetics, graphics, language, or control complexity that are unrelated to the challenge as purportedly captured by this questionnaire. This result should be considered as a form of correlation that accords with the conceptualisation of challenge in the questionnaire. It would take many more carefully controlled studies to provide firm evidence that the differences in the experience of challenge seen here are indeed due to the hypothesised challenges in the gameplay. However, a comprehensive analysis of the challenge structure in these specific games has not been our goal here.

9. Discussion

In this article, we have described the development of the *Challenge Originating from Recent Gameplay Interaction Scale* (CORGIS) and reported basic psychometric properties for the measure and preliminary construct validity. We filled in a need for a new measurement tool amongst existing questionnaires by developing and refining a standardised and reliable scale that can be used to assess perceived challenge as player experience in a range of video games for a variety of players. Our scale allows for an evaluation of performative, cognitive, emotional and decision-making challenge in video games.

9.1. Contributions

Research into challenge as player experience is still ongoing (Bopp et al., 2018), however, we hope that with the help of the new tool that measures perceived challenge in games, we can refine our understanding of this player experience. Our novel measurement instrument was developed based on the framework we built around existing literature and our interviews with players and games user researchers. Interestingly, but perhaps not surprisingly, this framework matches the findings of Bopp et al. (2018), providing additional confidence in our findings.

Being able to measure emotional challenge was one of the key aims of our research, as this concept has not received as much attention in existing empirical research as other more "traditional" cognitive and physical challenge types. The correlations between emotional challenge and other challenge types suggest that emotional challenge is indeed part of the experience of challenge in video games.

Another novel, although somewhat unexpected, discovery that came out of our studies was a new type of challenge, which has not yet been discussed in previous literature – decision-making challenge. In the existing literature, aspects of decision making have traditionally been subsumed under cognitive challenge, and our identification as a separate factor contributing to perceived challenge might thus come as surprise. However, both our qualitative and quantitative findings offer an explanation and thus have the potential to enrich the present conceptualisation of perceived challenge. Previously, game designers and researchers have either looked at one specific type of challenge, or compared the predominant types of cognitive and performative

challenge. But performative challenge is somewhat opposed to decision making, in that it relies on quick reactions rather than deliberate thinking. This is supported by statements from our interviews and the negative correlation in Table 6, and explains why decision making has previously been subsumed under cognitive challenge. By juxtaposing these traditional types with emotional challenge, our analysis has highlighted that decision making also plays a crucial role in the latter, e.g. when players must choose to sacrifice one character over another. This is also supported by the positive correlations Table 6. Rather than incorporating specific instances of decision-making for both cognitive and performative challenge, the CORGIS abstracts it into a separate factor which offers an overall more compact scale.

9.2. Limitations

The research presented in this paper should be interpreted in light of several limitations. First, scale development and validation is an ongoing process. While we have conducted several studies with almost 1400 players in total, more studies are needed to evaluate how the CORGIS performs across different genres, game modes and difficulty settings, and game audiences.

Second, our aim was to create an instrument comprehensive enough to capture experiences of players with varying levels of gaming experience and skill, while also being specific enough to describe different aspects of challenging experiences in different kinds of games. We believe that the CORGIS allows to do so. Nonetheless, no claims can be made with regards to the completeness of this tool. As work on challenge is ongoing (Bopp et al., 2018), we might learn more about this experience and the tool would need to be adjusted accordingly. However, to present day, this tool measures what we think challenge is.

We developed the CORGIS based on the experiences of players playing a range of different games from a variety of different genres. However, as more games get released every day, a possibility remains that some experiences created through different gameplay might not be fully captured by the CORGIS. Hence, future research will focus on testing the questionnaire on a range of games that were not initially considered in the literature review, our interviews or included in Bopp et al. (2018)'s survey.

10. Conclusion and Future Work

Challenge is an experience that many players seek in digital games, and it is central to the overall enjoyment of many games. In order to investigate how perceived challenge is formed and to be able to relate it to other player experiences, we as researchers need a reliable tool to measure the experience of challenge in video games. At the same time, such a measurement tool can support game developers in evaluating the challenge structure in their games more accurately but also more extensively, ultimately allowing for the design of more diversified and richer player experience.

Motivated by these needs, we developed and validated the *Challenge Originating from Recent Gameplay Interaction Scale* (CORGIS) as a tool for games research, design and testing. We have conducted a review of related work across games user research, game design and game AI. This allowed us to discriminate challenges or difficulties as intrinsic properties of games, and *perceived challenge* as the player experience that we would like to measure. We complemented and deepened this review by an open coding analysis of interview data with video game players and researchers. This was followed by a quantitative analysis of two large online surveys employing a larger and reduced version of our questionnaire respectively, covering the responses of 1400 players of varied skill on a wide range of games. As a result of these extensive, iterative studies, we have operationalised perceived challenge into four distinct factors: **Performative**, **Cognitive**, **Emotional**, and **Decision-Making Challenge**. Our quantitative results demonstrate that the developed questionnaire is a reliable and valid measure of challenge as player experience.

In particular, we have validated the questionnaire on three games with very different challenge structures. As a next step, we must investigate how the questionnaire performs for a single game with different difficulty modes and players with different skill levels and backgrounds. This would allow us to test the sensitivity and discriminant validity of the questionnaire. To further probe into the experience of challenge in video games, additional empirical studies need to be conducted to explore the relationship between perceived challenge and other player experiences, including immersion, uncertainty, autonomy and competence, and emotional experiences; as well as relating perceived challenge to objective data about players' performance. Our extensive literature review and interview studies presented here allows formulating preliminary hypotheses about these relationships.

We are confident that this scale is an essential tool that will assist researchers and game developers in learning more about perceived challenge and its relationship to player enjoyment, ultimately leading to better games.

References

Abuhamdeh, S., Csikszentmihalyi, M., 2012. The importance of challenge for the enjoyment of intrinsically motivated, goal-directed activities. Personality and Social Psychology Bulletin 38, 317–330. URL: https://doi.org/10.1177/0146167211427147, doi:10.1177/0146167211427147.

Abuhamdeh, S., Csikszentmihalyi, M., Jalal, B., 2015. Enjoying the possibility of defeat: Outcome uncertainty, suspense, and intrinsic motivation. Motivation and Emotion 39, 1–10.

Adams, E., 2014. Fundamentals of game design. Pearson Education.

Bateman, S., Doucette, A., Xiao, R., Gutwin, C., Mandryk, R.L., Cockburn, A., 2011. Effects of view, input device, and track width on video game driving, in: Proceedings of Graphics Interface 2011, Canadian Human-Computer Communications Society. pp. 207–214. URL: http://dl.acm.org/citation.cfm?id=1992917.1992952.

Blandford, A., 2013. Semi-structured qualitative studies, in: Soegaard, M., Dam, R.F. (Eds.), The Encyclopedia of Human-Computer Interaction, 2nd Ed.. Interaction Design Foundation.

Bontrager, P., Khalifa, A., Mendes, A., Togelius, J., 2016. Matching Games and Algorithms for General Video Game Playing, in: Proceedings of the 12th International Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE).

Bopp, J.A., Mekler, E.D., Opwis, K., 2016. Negative emotion, positive experience?: emotionally moving moments in digital games, in: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, ACM. pp. 2996–3006.

Bopp, J.A., Opwis, K., Mekler, E.D., 2018. "an odd kind of pleasure": Differentiating emotional challenge in digital games, in: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, ACM. p. 41.

Braun, V., Clarke, V., 2006. Using thematic analysis in psychology. Qualitative research in psychology 3, 77–101.

Cairns, P., 2019. Doing Better Statistics in Human-computer Interaction. Cambridge University Press.

Chalmers, R.P., et al., 2012. mirt: A multidimensional item response theory package for the r environment. Journal of Statistical Software 48, 1–29.

Charmaz, K., 2014. Constructing grounded theory. Sage.

Chen, J., 2007. Flow in games (and everything else). Communications of the ACM 50, 31–34.

Cole, T., Cairns, P., Gillies, M., 2015. Emotional and functional challenge in core and avant-garde games, in: Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play, ACM. pp. 121–126.

Cox, A., Cairns, P., Shah, P., Carroll, M., 2012. Not doing but thinking: the role of challenge in the gaming experience, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM. pp. 79–88.

Csikszentmihalyi, M., 1990. The psychology of optimal experience. Harper & Row, New York.

Davis, J.P., Steury, K., Pagulayan, R., 2005. A survey method for assessing perceptions of a game: The consumer playtest in game design. Game Studies 5.

De Kort, Y.A., IJsselsteijn, W.A., Poels, K., 2007. Digital games as social presence technology: Development of the social presence in gaming questionnaire (spgq). Proceedings of PRESENCE .

Deci, E.L., Ryan, R.M., 2000. The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. Psychological inquiry 11, 227–268. URL: `www.jstor.org/stable/1449618`.

Denisova, A., Cairns, P., 2015. The placebo effect in digital games: Phantom perception of adaptive artificial intelligence, in: Proceedings of the 2015 annual symposium on computer-human interaction in play, ACM. pp. 23–33.

Denisova, A., Guckelsberger, C., Zendle, D., 2017. Challenge in digital games: Towards developing a measurement tool, in: Proceedings of the 2017 CHI

Conference Extended Abstracts on Human Factors in Computing Systems, ACM. pp. 2511–2519.

Dunn, T.J., Baguley, T., Brunsden, V., 2014. From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. British journal of psychology 105, 399–412.

Embretson, S.E., Reise, S.P., 2000. Item response theory for psychologists. Psychology Press.

Ermi, L., Mäyrä, F., 2005. Fundamental components of the gameplay experience: Analysing immersion. Worlds in play: International perspectives on digital games research 37, 2.

Feil, J., Scattergood, M., 2005. Beginning game level design. Thomson Course Technology.

Fornell, C., Larcker, D.F., 1981. Evaluating structural equation models with unobservable variables and measurement error. Journal of marketing research 18, 39–50.

Gee, J.P., 2008. Learning and games. The ecology of games: Connecting youth, games, and learning 3, 21–40.

Hair, J.F., Anderson, R.E., Tatham, R.L., Black, W.C., 1998. Multivariate data analysis. 1998. Upper Saddle River .

Horn, J.L., 1965. A rationale and test for the number of factors in factor analysis, in: Psychometrika, Springer-Verlag. pp. 179–185. URL: https://doi.org/10.1007/BF02289447, doi:10.1007/BF02289447.

Hudson, M., Cairns, P., 2014. Measuring social presence in team-based digital games. Interacting with Presence: HCI and the Sense of Presence in Computer-mediated Environments , 83.

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, ACM. pp. 429–433.

Hunicke, R., LeBlanc, M., Zubek, R., 2004. Mda: A formal approach to game design and game research, in: Proceedings of the AAAI Workshop on Challenges in Game AI, p. 1.

Isbister, K., 2016. How games move us: Emotion by design. Mit Press.

Jennett, C., Cox, A.L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., Walton, A., 2008. Measuring and defining the experience of immersion in games. International journal of human-computer studies 66, 641–661.

Juul, J., 2009. Fear of failing? the many meanings of difficulty in video games. The video game theory reader 2, 237–252.

Klarkowski, M., Johnson, D., Wyeth, P., McEwan, M., Phillips, C., Smith, S., 2016. Operationalising and evaluating sub-optimal and optimal play experiences through challenge-skill manipulation, in: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, ACM. pp. 5583–5594.

Klimmt, C., 2003. Dimensions and determinants of the enjoyment of playing digital games: A three-level model, in: Level up: Digital games research conference, pp. 246–257.

Klimmt, C., Blake, C., Hefner, D., Vorderer, P., Roth, C., 2009. Player performance, satisfaction, and video game enjoyment, in: International Conference on Entertainment Computing, Springer. pp. 1–12.

Klimmt, C., Hartmann, T., Frey, A., 2007. Effectance and control as determinants of video game enjoyment. Cyberpsychology & behavior 10, 845–848.

Kline, P., 2014. An easy guide to factor analysis. Routledge.

Lazzaro, N., 2004. Why we play games: Four keys to more emotion without story .

Loftus, G.R., Loftus, E.F., 1983. Mind at play; The psychology of video games. Basic Books, Inc.

Lomas, J.D., Koedinger, K., Patel, N., Shodhan, S., Poonwala, N., Forlizzi, J.L., 2017. Is difficulty overrated?: The effects of choice, novelty and suspense on intrinsic motivation in educational games, in: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, ACM. pp. 1028–1039.

Malone, T.W., 1981. Toward a theory of intrinsically motivating instruction. Cognitive science 5, 333–369.

Malone, T.W., 1982. Heuristics for designing enjoyable user interfaces: Lessons from computer games, in: Proceedings of the 1982 conference on Human factors in computing systems, ACM. pp. 63–68.

Malterud, K., Siersma, V.D., Guassora, A.D., 2016. Sample size in qualitative interview studies: guided by information power. Qualitative health research 26, 1753–1760.

Nielsen, T.S., Barros, G.A.B., Togelius, J., Nelson, M.J., 2015. General Video Game Evaluation Using Relative Algorithm Performance Profiles, in: Proceedings of the 18th Conference on Applications of Evolutionary Computation, Springer. pp. 369–380.

Nunally, J.C., Bernstein, I.H., 1978. Psychometric theory.

Petralito, S., Brühlmann, F., Iten, G., Mekler, E.D., Opwis, K., 2017. A good reason to die: How avatar death and high challenges enable positive experiences, in: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, ACM. pp. 5087–5097.

Poels, K., De Kort, Y., Ijsselsteijn, W., 2007. It is always a lot of fun!: exploring dimensions of digital game experience using focus group methodology, in: Proceedings of the 2007 conference on Future Play, ACM. pp. 83–89.

Power, C., Cairns, P., Denisova, A., Papaioannou, T., Gultrom, R., 2018. Lost at the edge of uncertainty: Measuring player uncertainty in digital games. International Journal of Human–Computer Interaction , 1–13.

Reise, S.P., 2012. The rediscovery of bifactor measurement models. Multivariate behavioral research 47, 667–696.

Rouse III, R., 2010. Game design: Theory and practice. Jones & Bartlett Learning.

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

Salen, K., Zimmerman, E., 2004. Rules of play: Game design fundamentals. MIT Press.

Salisbury, J.H., Cole, T., 2016. Grounded theory in games research: Making the case and exploring the options, in: Proceedings of the 1st International Joint Conference of DiGRA and FDG.

Schell, J., 2014. The Art of Game Design: A book of lenses. CRC Press.

Shneiderman, B., 1982. Direct manipulation: A step beyond programming languages. ACM SIGSOC Bulletin 13, 143.

Sicart, M., 2014. Play matters. MIT Press.

Thomas, M.L., 2011. The value of item response theory in clinical assessment: a review. Assessment 18, 291–307.

Vorderer, P., Hartmann, T., Klimmt, C., 2003. Explaining the enjoyment of playing video games: the role of competition, in: Proceedings of the Second International Conference on Entertainment Computing, Carnegie Mellon University. pp. 1–9.

Yannakakis, G.N., Hallam, J., 2007. Towards optimizing entertainment in computer games. Applied Artificial Intelligence 21, 933–971.

Yannakakis, G.N., Hallam, J., 2009. Real-time game adaptation for optimizing player satisfaction. IEEE Transactions on Computational Intelligence and AI in Games 1, 121–133.

Appendix: Using the CORGIS

The CORGIS is a psychometric instrument for measuring perceived challenge in video games. The questionnaire allows to measure four types of perceived challenge in games: cognitive, performative, emotional, and decision-making challenge – four subscales of the CORGIS:

Performative Challenge: arising from the game requiring rapid and accurate action from the player.

Cognitive Challenge: arising from the need for planning ahead, memorisation, effort, preparation and multi-tasking.

Emotional Challenge: arising from the emotions evoked in the player which might also have implications for things they thought about outside of the game.

Decision-Making Challenge: arising from having to make choices that were difficult or could lead to regrettable outcomes.

EVALUATING PERCEIVED CHALLENGE IN A GAME

The questionnaire is meant to be administered immediately after the game session has finished.

Instructions for players

Please indicate to what extent you agree/disagree with each of the statements in Table 9 after you finished playing the game on the following scale:

☐ (1) Strongly Disagree
☐ (2) Disagree
☐ (3) Slightly Disagree
☐ (4) Neither Agree Nor Disagree
☐ (5) Slightly Agree
☐ (6) Agree
☐ (7) Strongly Agree

The scores for each subscale are then computed as the sum value of its items ranked on the scale from 1 to 7, as indicated above.

Subscales	Items
Cognitive Challenge	CC1. Succeeding in the game required much planning
	CC2. I had to memorise a lot of different things when playing the game
	CC3. I had to think several steps ahead when playing the game
	CC4. I had to prepare for the things that the game threw at me
	CC5. Playing the game requires great effort
	CC6. I felt challenged when playing the game
	CC7. I had lots of different things to think about at once in the game
	CC8. The game made me manage several tasks at the same time
	CC9. I had to constantly keep track of what was going on in the game
	CC10. I had to think actively when playing the game
	CC11. Playing the game required me to do my best
Emotional Challenge	EC1. This game is more than just a game to me
	EC2. The things that happened in the game made me sad
	EC3. I invested much thought into the game
	EC4. I felt a sense of responsibility for characters and events in the game
	EC5. The game made me think about real life issues
	EC6. Playing the game was stimulating
	EC7. I felt a sense of suspense when playing the game
	EC8. The game had moral dilemmas in it where the choice was not obvious
	EC9. The game involved making moral choices that I didn't agree with
Performative Challenge	PC1. I had to react quickly when playing the game
	PC2. I had to act quickly when playing the game
	PC3. Thinking fast was an important part of the game
	PC4. Quickly responding to things that I saw was an important part of the game
	PC5. I had to make snap decisions when playing the game
Decision Making Challenge	DMC1. There were some decisions in the game that I regretted
	DMC2. I wonder how different the outcome in the game would be had I chosen a different option
	DMC3. I had to make difficult choices in the game
	DMC4. I had to think about possible alternatives for my actions in the game
	DMC5. The game made me think hard about my decisions

Table 9: Four subscales and the corresponding items of the Challenge Originating from Recent Gameplay Interaction Scale (CORGIS).