User-adapted Gamification: developing a user-centred design feature preference model to inform engaging design

By

HARKEERET SINGH LALLY

Supervised by Dr. Pelham Carter and Dr. Kyle Brown

A thesis submitted to the Faculty of Business, Law and Social Sciences

Of Birmingham City University

For the degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

Faculty of BLSS

Birmingham City University

Submitted: 11/10/2022

Acknowledgements

Having worked towards the submission of my PhD for just over five years, at the time of writing these acknowledgements, the decision of which colleagues, friends, or family, ought to be mentioned alongside this small and humble, yet personally persisting body of work, requires nothing more than a moment of thought.

First and foremost are my two PhD supervisors, both of whom have supported and mentored my academic development prior to and throughout the PhD. To Kyle and Pelham, thank you so much for the guidance and wisdom you share, your endless feedback, your precious time, and mostly, the opportunity to know you personally. Commanding academics in your own right, I have often felt a strong sense of security and assurance that you are both in my corner, and with this comes an almost irresponsible level of confidence.

Second, my older brother, Marty. Were it not for your influence, I believe many of my interests, perspectives, and the attitudes I hold towards the world around me might never had developed in the way they have. You will most likely never read this, and therefore I will say I continue to look up to you and your ability but will display a complete lack of interest in anything you have to say.

Third, to my family, which has grown much larger since becoming a husband. To my parents Prem and Kam, thank you for sacrificing so much throughout your life so that I could pursue opportunities I would otherwise be unable to. Your efforts have and never will go unnoticed. To my second Father Ranjit, thank you for providing me with my first and favourite working environment. Your gift of opportunity introduced me to the benefits of being responsible, accountable, and how to lead.

Finally, to Henna. You were my crush at secondary school, my girlfriend during sixth form and undergrad, my fiancé during our PhDs, and now you are my wife. In a few years I will have lived longer being in love with you, than the years I spent not knowing you. You have nurtured my vulnerabilities, encouraged my adventure, supported my pursuits, and always demonstrated an uncompromising belief in my ability. To me, there exists no medium with which I could comprehensively represent the deeply profound impact you have had on my development as a person, nor the way in which you have wholly improved the trajectory of my life, and the life of our future children. A true soulmate.

Table of Contents

Abstract

Gamification is a behavioural intervention that applies game-like elements to non-game contexts (Deterding et al., 2011) for the purpose of increasing performance of a target behaviour within a non-game context (or task therein). Existing literature highlights a substantial number of instances wherein Gamification is unsuccessful, such that applied design features elicit little to no impact on a target behaviour. The field of Adaptive Gamification seeks to improve the effectiveness of Gamification, by adopting a user-centred design approach wherein the design features used to increase the performance of a target behaviour are tailored or "adapted" to meet an often-unique set of user needs.

Existing methodologies which support the Adaptive Gamification approach are, however, limited. Principally, there exists no model which can effectively measure the level of preference an end user possesses towards a given design feature. In the context of how research can inform Adaptive Gamification design, understanding the level of preference a user possesses towards a given design feature is of critical importance, given that this relationship can directly inform design of a user-centred and tailored Gamification experience. This doctoral research project sought to develop a design feature preference model which could be used to accurately capture the design feature preference of users and provide insight into which design features users are likely to be more receptive to. To this end, the doctoral research project aimed to fulfil three research aims.

The first research aim was to develop a model which could measure user design feature preference, the fulfilment of was achieved across Study one, Study two, and Study three. Combined across all three studies, data from 2322 players was analysed. The first of these studies (Study one) operationalised a total of 37 design features (later increased to 47) using vignette methodology to describe the functionality and purpose of each feature. The second of these studies (Study two) subjected the 47 design features to an Exploratory Factor Analysis (EFA) that returned a nine-factor solution. The third of these studies (Study three) furthered model development, by subjecting the nine-factor solution to a Confirmatory Factor Analysis (CFA), which confirmed the nine-factor solution (though some amendments and reallocation of items were made).

The second research aim was to identify how user characteristics could predict design feature preference, the fulfilment of which was achieved across Study two and Study

three. Combined across both studies, data from 2011 players was analysed. The first of these studies (Study two) measured user design feature preference, user motivation, and user personality. The results of this study revealed which design features were most likely to predict variance in user engagement, as well as which motivations and personality traits were associated with predicting variance in preference for these design features. Using the same methodology, the second of these studies (Study three) expanded the range of user characteristics measured in relation to design feature preference, by measuring Gamefulness (a concept stemming from the area of Gamification that refers to what aspects of a gaming experience the end user values). Study three also revealed how Gamefulness could impact design feature preference.

The third research aim was to test whether any relationships between user characteristics and design feature preference would correspond to tangible difference in user engagement, the fulfilment of which was achieved across Study four and Study five. Combined across both studies, data from 96 players was analysed. The first of these studies (Study four) sought to test the relationships between design feature preference and user engagement, when measuring user engagement via an online task-performance experiment, wherein participants were asked to play a series of online browser games (selected due to the design features they comprised of) while their engagement was measured. The second of these studies (Study five) sought to improve on the measurement of user engagement using in-game behavioural metrics, which is argued as a more representative account of user engagement. Both studies returned non-significant results, which were not consistent with relationships identified in Study two and Study three, though the role of methodological limitations in these findings are extensively discussed at the end of each study chapter.

1. Literature review

1.1 History of play and modern gaming

Play refers to a range of voluntary, intrinsic, and recreationally motivated activities one performs to derive enjoyment and mitigate boredom (Garvey, 1990). Play behaviour is well recorded, with some of the earliest games such as Knucklebones and the Royal Game of Ur being played between 4000-6000 years ago (David, 1998), across multiple regions, such as the Mediterranean, the Middle East (Shenk, 2011), China, and Europe (Sidéra & Vornicu, 2016).

The technological revolutions of the 20th century facilitated the development of a new form of play that changed the way in which play behaviours are conceptualised, observed, and understood today. The emergence of the digital computer which was eventually adopted by universities, government bodies, and corporations, prompted the development of the first digital computer games (Kent, 2010). From 1950-1970, computer technology would significantly improve, eventually leading to the commercialisation of video games, with the years that followed being characterised by a rapidly evolving video game market. For instance, the arcade game industry experienced a steep rise in popularity from 1978-1982, with revenues from coin-operated games jumping from $308 million to $7.7 billion (Larsen & Rogers, 1984). Continuous development in the industry has led to video games today being more technologically sophisticated, more diverse, more accessible, such that there will be over 3.31 billion worldwide players by 2024 (Yanev, 2022). At its height in popularity, video games are now played by most age groups and genders, (Kahn et al., 2015), and are considered a key medium of recreation (McGonical 2011) with the average household having at least one technology solely dedicated to video gaming (Entertainment Software Association, 2015). In the context of the 2020 pandemic, gaming activity increased substantially, further highlighting the integrity of gaming as forming a part of many daily routines (Barr & Copeland-Steward, 2021).

1.2 Gaming and research

Current gaming literature highlights the role of user psychology in influencing player behaviours and preferences, examples of which can be found with the areas of user motivation (Kardefelt-Winther, 2014), and user personality (Braun et al., 2016). For example, the motivation a user fosters when playing videos games can mediate the games they prefer to play (O'Brien, Gagnon, Egan, & Coulter, 2022) as well as their overall game activity, such as how long they spend playing (Hoffman & Nadelson, 2010) and what they do when they play (Billieux et al., 2013). Additionally, several studies highlight the traits which comprise a user's overall personality can also influence their game preference (Zammitto, 2010; Quick et al., 2012) and in-game behaviours (Worth & Book, 2014).

Gaming has also been considered within clinical contexts (Kuss, 2013), with particular focus on video gaming's association with violence, addiction and depression (Granic et al., 2014). More recent developments supporting the clinical recognition of adverse video gaming outcomes include the DSM-V classifying internet gaming addiction as an area warranting further study (Petry et al., 2015) and in its 11th revision of the International Classification for Diseases, the World Health Organisation recognised problematic video gaming behaviour as a mental health disorder (Aarseth, 2017). Though there is a general research emphasis on the problematic aspects of video gaming, Ferguson and Colwell (2017) argue that this could be attributed to generational conflicts and a lack of video game understanding and experience, characterised by paradigms which emerged when gaming was not as sophisticated as it is today. Moreover, polarised perspectives on gaming, such as being characteristically socially isolative, have been debunked by studies that highlight the wealth of social opportunities provided by modern games (Kaye et al., 2018).

Adopting an alternative perspective, research has also focused on the beneficial uses of video gaming (Cade & Gates, 2017). Granic et al. (2014) suggest that video games can provide a context within which cognitive, motivational, emotional and social benefits can be experienced. For instance, improved visuospatial cognition (Castel et al., 2005; Green & Bavelier, 2003; Green & Bavelier, 2007; Rosser et al., 2007), post-gaming positive mood changes (Fleming & Rick-Wood, 2001), and likelihood of performing prosocial behaviour (Ewoldsen et al., 2012; Greitemeyer et al., 2010). Building on the notion that the positive outcomes created by gaming are not limited or exclusive to the systems or services from

which they originate, researchers have also explored how the principles of gaming could be extrapolated to real-life settings. These positive impacts can be leveraged in a way that can lead to improved productivity (Arai et al., 2014), well-being (Jones et al., 2014), and prosocial behaviour (Bang et al., 2007). Central to this discussion is the concept of Gamification (Sailer et al., 2017), which is a behavioural intervention that applies game-like elements to non-game contexts (Deterding et al., 2011) for the purpose of increasing performance of a target behaviour within a non-game context (or task therein). Some common examples of design features include Points, Badges, and Leaderboards (Codish & Ravid, 2014), each of which serve different functions within a gamified context, but principally aim to increase user engagement. Table 1 details the functions of these design features and shows how they appear in the original video game context (further illustrating how they might be applied in a gamified context).

Table 1

Functions of design features

Design feature	**Function**	**Appearance**	**Game**
Points	A measurement which quantifies progress made or level of skill achieved by a user		New Super Mario Bros. U (2012, Nintendo Wii)

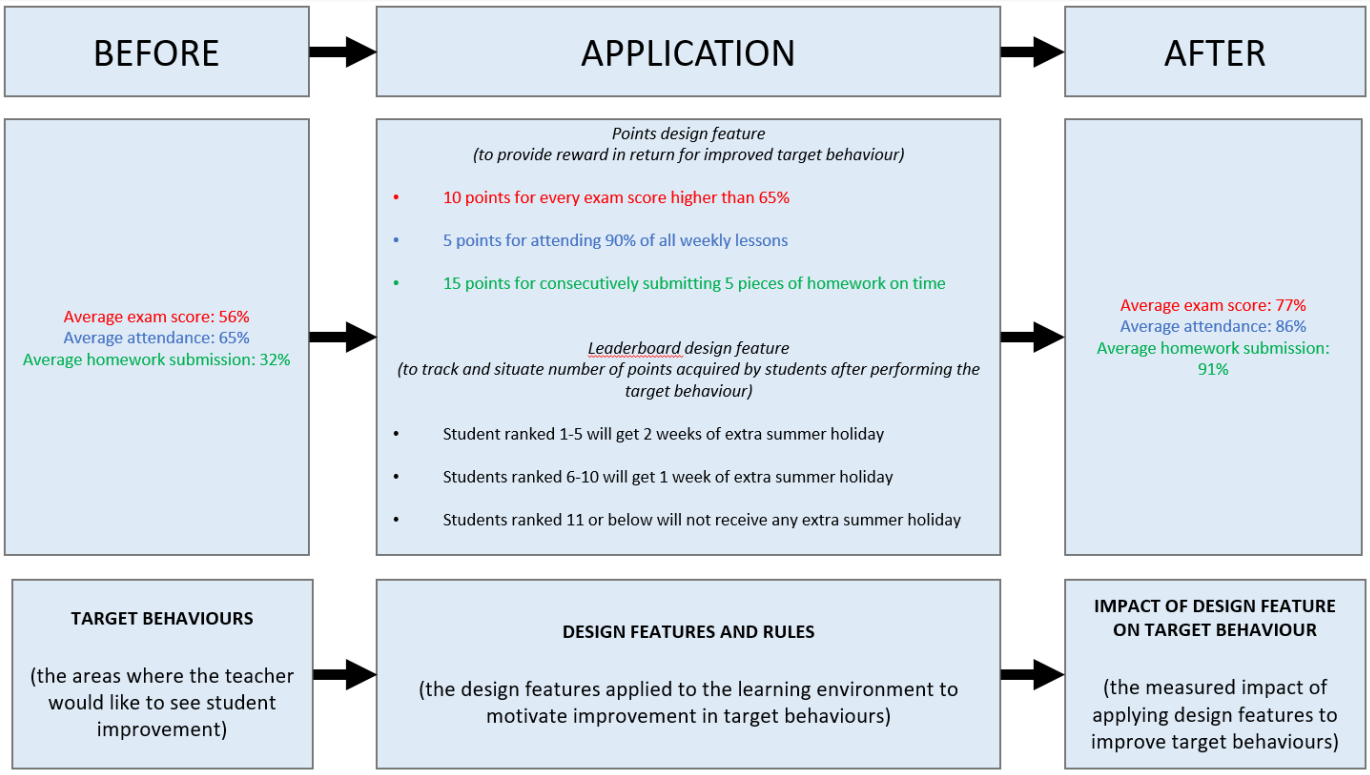
Badge	A (virtual) accessory which typically represents a user's achievements, accolade, or membership		Call of Duty: Advanced Warfare (PC, 2014)
Leaderboard	A scoreboard which situates a player's progress or skill with peers, often organised in order of high to low		Gears of War: Ultimate Edition (XBOX, 2015)

Hamari et al. (2014) conceptualise the Gamification process as comprising of three parts: the first part concerns the implementation of design features; the second is the way in which interaction with the implemented design features impacts or targets aspects of user psychology; and the third concerns the way in which impacts on user psychology by way of interaction with the implemented design features, can drive differences in the target behaviour.

A conceptual flow diagram detailing the process of Gamification is illustrated in Figure 1, taking the example of a teacher who wishes to gamify student learning. The teacher applies design features to the classroom, and in doing so incentivises students to acquire a higher rank and receive the reward. As the rules indicate that increasing rank is achieved by completing target behaviours, students will perform the target behaviours,

and indirectly improve their learning via repeat exposure and engagement when pursuing a higher rank to receive the reward.

Figure 1 - Gamification process flow diagram



1.3 Gamification: current perspectives and problems

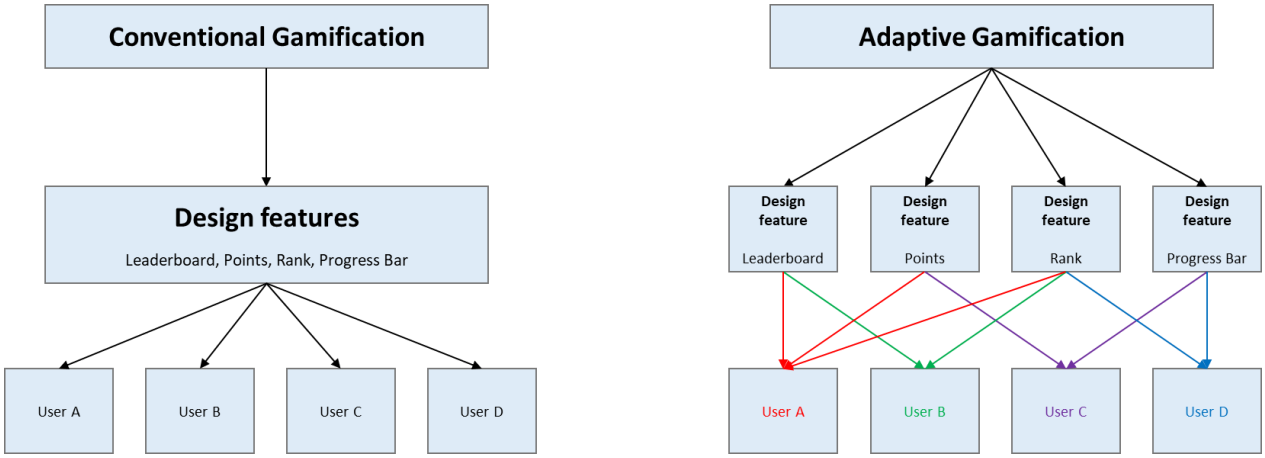
Evidence for the effectiveness of Gamification is well recorded, with studies highlighting success in improving workplace productivity (where software technicians were 150% more productive in identifying and removing bugs following the implementation of a Points design feature; Arai et al., 2014), student learning (where students who were presented with gamified tasks engaged with their project 30 times more frequently than students who were not presented with gamified tasks following the implementation of a Leaderboard design feature; Landers & Landers, 2014); and increased health behaviours (where children who were presented with a storyline increased daily vegetable consumption following the implementation of a Levels design feature; Jones et al., 2014).

By contrast, there are examples where Gamification fails to change a target behaviour and is ineffective, such as in a learning context (where the reward of Badges for completing tasks did not result in any significant differences in student learning between a control and treatment group; Hakulinen et al., 2013) or in an occupational context (where

window dressing aspects of games, such as Points or Leaderboards were indiscriminately applied to a range of occupational contexts but did not produce meaningful changes in target behaviours; Landers, 2015). According to Lopez and Tucker (2019), cases of Gamification not being successful are often the result of researchers employing a one-size-fits-all approach (Nacke & Deterding, 2017), which presupposes that users of a gamified system will react and respond the same when interacting with a given design feature (Jia et al., 2016; Klock et al., 2015; Lopez & Tucker, 2019). Users of gamified systems, however, appear to perceive, react, and respond differently when interacting with different design features (Hamari & Tuunanen, 2014). As such, the subsequent effect of Gamification on driving a target behaviour could vary significantly between participants (Barata et al., 2017; Fitz-Walter et al., 2017). For example, the subsequent outcomes following interaction with a given design feature might improve one user's performance and overall engagement, but in contrast there could be a more negative impact on performance and level of engagement for another participant (Witt et al., 2011).

The recognition that users are likely to respond differently when interacting with design features has been referred to as Adaptive Gamification, which can be considered as a type of Gamification that seeks to adapt or tailor design to meet the needs of the user (Böckle et al., 2017). Figure 2 provides a visual illustration of how conventional and Adaptive Gamification differ in their fundamental approach to design. Conventional Gamification sees the uniform application of four design features, such that all users encounter and interact with all design features during their user experience. By contrast, Adaptive Gamification sees a more tailored and user-specific application of the design features, such that only some design features are interacted with and encountered by some users. For example, in the conventional Gamification design, User C will interact with all four design features. However, in the Adaptive Gamification design, User C will only interact with two design features (Points and Progress Bar).

Figure 2 - Difference in design approaches of conventional and Adaptive Gamification



The Adaptive Gamification approach draws its ontological principles from the field of differential psychology; the psychological study of how (and why) variance in human behaviour occurs (Stern, 1900), with the individual's personal characteristics constituting the core variables of measurement for psychological study (Wilpert, 2001). A variety of dimensions (otherwise referred to as individual differences) in which substantial subject variability occurs, have been extensively researched. Two of these areas are user motivation (Pluymen et al., 2021) and user personality (Silvia & Christensen, 2020). In the context of Adaptive Gamification, individual differences are of central importance, given that the personal characteristics of the user are thought to have a causal or moderating impact on how effective the design of Gamification can be.

For example, user motivation has been found to predict which game genre a user is likely to prefer. For example, Ghuman and Griffiths (2012) found that those with those who are motivated to achieve tending to prefer first-person shooters over real-time strategy, while Bijvank et al. (2012) found user engagement and play behaviour varied between participants who held different play motivations. In this context, the user plays a video game to fulfil a fostered motivation. Findings also reveal how personality might mediate design feature preference. For instance, Jia et al. (2016) report that extraverts are particularly motivated by Points, Levels, and Leaderboard design features. Additionally, higher levels of openness are found to correspond to greater preference for feedback design features (Denden et al., 2017), while higher levels of extraversion correspond to greater preference for design features which enable communication with others (Daughenbaugh et al., 2002). In this context, the user plays a video game that

complements existing trait tendencies. Taken together, research demonstrates the influence of individual differences on user preference, specifically for the genres users enjoy and the design features they prefer. As such, in the interest of maximising user engagement, researchers may benefit from leveraging the psychological insights related to individual differences, such that a more tailored form of Gamification which complements, or targets user preferences could be designed.

1.4 Operational barriers to Adaptive Gamification

Despite initial research insights into how Gamification design can be adapted to maximise user engagement, there exists an operational barrier which limits a more nuanced and prescriptive understanding of how research can inform Gamification prospects. To illustrate, it can be argued that to understand precisely how the design of Gamification can be tailored to maximise user engagement, user preference for design features must first be operationalised, such that variance in user individual differences (e.g., motivation or personality) can be assessed in relation to user preference for a given design feature. Such a prerequisite would provide researchers with a start point (individual difference) and end point (preference for design feature), between which any relationship could be identified, scrutinised, tested, and most critically, leveraged to inform more effective Gamification design. Given the role of user needs in determining how long and how much they interact with a given system, identifying this in the context of Gamification could lead to more effective Gamification design, and consequently improvement in a target behaviour (e.g., exercise or learning).

For example, if researchers were to examine how Gamification can be effectively adapted to meet the needs of users who score high in the personality trait of conscientiousness, there exists two key variables in need of measurement. The first, would be the user's level of trait conscientiousness, while the second would be the users reported level of preference for a design feature of interest.

At present, a review of existing Gamification literature indicates that, although there are several examples of operationalised start points, with scales that measure a broad-spectrum of individual differences, such as (Kahn et al., 2015) and personality (Ashton & Lee, 2004), there are no existing operationalised measures that would capture user design feature preference (preference for a given design feature), thus researchers do not

possess a distinct end point against which variance in individual differences can be paired or assessed against. The absence of an operationalised measure that captures distinct user design feature preference might explain why the paradigm has only recently emerged (Böckle et al., 2017).

In developing a measure that can capture user design feature preference, additional issues present in Gamification literature can also be resolved, namely the scope of design features which are researched. For example, Nacke and Deterding (2017) note that most Gamification research focuses on the utility of three main design features: Points, Badges, and Leaderboards. The excessive focus on these design features could be attributed to a relative ease of implementation (these design features in particular are not difficult to implement into a non-game context) or the more observable profile they possess (they are more tangible and directly visible to a user) (da Rocha Seixas et al., 2016). In operationalising design features, such that associated user preference can be measured, the number of design features involved in Gamification research will inevitably increase. Such an outcome is of particular importance to Adaptive Gamification, given the emphasis on diversifying the design of Gamification to meet the expanding and varied needs of a technologically evolving user base (otherwise referred to as digital natives) (Högberg et al., 2019).

Moreover, by expanding the range of design features available to research or implement in Gamification design, a broader issue of meeting the evolving needs of digital users might also be resolved. According to Högberg et al. (2019), as the presence of game principles and design features continue to suffuse into normal aspects of life, the existing and coming generation of technology users are likely to possess significant familiarity with a variety of design features. In the absence of novelty effects (the positive spike in performance or engagement following interaction with something new; Jeno et al., 2018), research indicates that engagement can be difficult to maintain within gaming contexts (Merikivi et al., 2017). As such, the diversification of design features used within research could provide a solution to the potential problem of increased user familiarity and meet the rapidly evolving characteristics and needs of a more technological user base.

1.5 Summary

Gamification is reported as an effective strategy to increase productivity and performance across a variety of contexts. Despite these positive findings, however, there are numerous instances where it is ineffective, which in most cases can be attributed to researchers approaching Gamification design from a one-size-fits-all perspective (Lopez & Tucker, 2019). The Adaptive Gamification approach provides an alternative perspective that seeks to improve on the limitations of the one-size-fits-all viewpoint, such that Gamification is designed by leveraging the psychological insights from the individual needs of each user to tailor and curate a highly engaging and more receptive user experience. Part of this process is the identification of user needs, which once recognised must be operationally paired or matched with design features of interest, thus providing prescriptions on how Gamification can be practically designed to meet any identified user needs. At present, there exists many psychometrically validated measures that capture how users score across a variety of individual differences, such as user motivation or personality. However, there are currently no instruments which measure user preference for design features, thus there exists an operational barrier to identifying how the needs of users can be practically fulfilled by the design of Gamification. The case for why a measure must be developed comprises of a focus on increasing the scope of design features made available in research, to resolve the broader concern of how a lack of diversity in design can limit the overall longevity of Gamification as a behaviour modification strategy.

2. Thesis aims

It was expected that the PhD will make practical and theoretical advances in Adaptive Gamification, by enabling researchers to further understand the relationship between theoretical user needs (indicated by individual differences) and practical need fulfilment via Gamification design (indicated by design feature preference). A series of five research studies were conducted, each of which would address the component research aims necessary to fulfil the overarching PhD aim. A more detailed summary can be found below, accompanied by a supplementary wireframe illustration found in Figure 3.

2.1 Overarching aim of PhD

The overarching aim of the PhD was to identify how the needs of users can be fulfilled via the adaptive design of Gamification, central to which was the development of a research instrument that could operationally measure user design feature preference.

2.2 Component research aims

2.2.1 Research aim one (Study one)

Develop a measure to operationalise design feature preference

The first research study focused on developing a research instrument which could measure user design feature preference. The study employed an online survey methodology and focused on establishing which items participants found to be most descriptive and representative of a given design feature. The items which received highest ratings of representativeness were compiled into a single questionnaire, referred to as the Design Feature Preference Scale (DFPS).

2.2.2 Research aim two (Study two and three)

Identify how the individual differences of users could predict design feature preference

The second research study focused on applying the DFPS to measure user preference for a selection of design features. Along with this, the second research study also collected individual differences data for user motivation and personality, to identify how design feature preference would relate to the personal characteristics of the user. Using the data,

the second research study also included further model development of the DFPS, with the inclusion of an EFA. The study employed an online survey methodology.

The third research study focused on further progressing the outcomes of study two. Repeating the study design and methodologies used in study two, the DFPS was applied to a new sample to measure user preference for a selection of design features. Additionally, a third individual difference of Gamefulness was applied to identify how design feature preference would relate to the personal characteristics of the user. The data from the third research study was also used to conduct CFA as part of further DFPS model development.

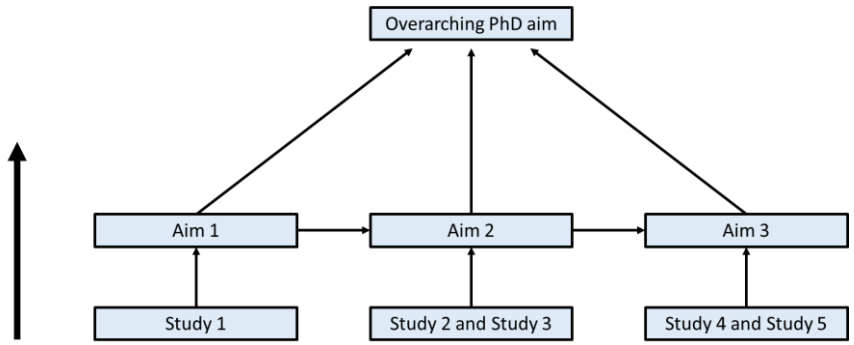
2.2.3 Research aim three (Study four and five)

Substantiate whether identified relationships would correspond to objective differences in user engagement.

The fourth research study focused on examining whether the previously identified relationships between user design feature preference and self-report measures of user engagement would be consistent with objective measurements of user engagement. The study employed an online experimental methodology that required participants to play a series of online games, during which their engagement was measured.

The fifth research study also focused on examining the previously identified relationships and whether objective variance of user engagement would be observed. However, unlike the fourth research study, the fifth research study used in-game behaviour metrics as a measurement of user engagement. The study employed an online survey methodology where participants were required to source and report a variety of in-game behavioural metrics from their game profile.

Figure 3 - Wireframe illustration of PhD research aims



3. Methodology

This chapter aims to comprehensively cover key aspects pertaining to the methodological approaches employed for the subsequent research studies discussed throughout the thesis. Broadly, this discussion comprises of six sections.

1. Study summary: an overview of the studies conducted throughout this thesis
2. Research philosophy: the epistemological and model of reasoning which underpin the broader research approach
3. Design strategy: the design strategy of the studies conducted
4. Data collection: the methods used to collect research data for the studies conducted
5. Sample selection: the reasoning behind sampling techniques for the studies conducted
6. Analysis techniques: the contextual reasoning for the analysis techniques used for the studies conducted

Throughout these sections, applicable reference is made to the studies in which discussion points are relevant.

3.1 Study summary

Study one focused on developing a research instrument which could measure user design feature preference, the outcome of which was the DFPS. Study two study focused on applying the DFPS to measure user preference for a selection of design features. Along with this, the second research study also collected individual differences data for user motivation and personality, to identify how design feature preference would relate to the personal characteristics of the user. Study three focused on further progressing the outcomes of study two. Repeating the study design and methodologies used in Study two, the DFPS was applied to a new sample to measure user preference for a selection of design features. Additionally, a third individual difference of Gamefulness was applied to identify how design feature preference would relate to the personal characteristics of the user. Study four focused on substantiating whether the previously identified relationships between user design feature preference and self-report measures of user

engagement would be consistent with a task-performance measurement of user engagement. Study five focused on assessing the predictive value of the DFPS, by examining whether user design feature preference would correlate with reported in-game behavioural metrics.

3.2 Research philosophy

3.2.1 Epistemology

The research conducted throughout this thesis has been approached using a positivist epistemological philosophy, which is characterised by a focus on identifying and generating explanatory associations and causal relationships (Xinping, 2002). By understanding how two given phenomena may be associated and how they may interact, there exists a starting foundation upon which these phenomena can be measured, controlled, and predicted (Alharahsheh & Pius, 2020). A positivist epistemological approach complemented the intended research outcomes of the PhD, such as identifying the mediating role of different individual differences (user motivation, personality, and Gamefulness) on user design feature preference. The importance of this relationship is discussed in more depth at the start of section 1.4. By understanding the associations between these two phenomena (individual difference and design feature preference), a user-centred approach to Gamification design would theoretically be more engaging for the end-user (due to an increased likelihood of user needs being fulfilled). Therefore, the relationship between two phenomena would be measured, controlled, and ultimately used to predict an intended outcome (improved user engagement with a gamified system).

3.2.2 Model of reasoning

Building on the positivist epistemological philosophy, the studies conducted throughout this PhD adopted the closely aligned hypothetico-deductive model of reasoning (Park et al., 2020), which proposes that scientific inquiry generally follow four stages: The first is to observe existing literature, the second is to formulate a prediction/hypothesis based on this literature, the third is to test the prediction/hypothesis, and the fourth is to evaluate whether testing confirms or rejects the hypothesis (Wicherts, 2017). An adherence to the hypothetico-deductive model of reasoning can be observed throughout

the studies conducted in this PhD. For example, based on observations from existing research, Study two and Study three posited that the individual differences of user motivation, personality, and Gamefulness, are likely to have a mediating role on user design feature preference. These predictions were then tested using the appropriate quantitative methods, after which the data was examined to assess if and how the prediction was supported or rejected.

3.3 Design strategy

All studies were conducted remotely and online. The main benefits of conducting the studies online include cost-effectiveness, with there being no fees incurred to design the surveys or post them across the range of Reddits used (though to incentivise participation, a voluntary cost was incurred for a participant-only game giveaway). Additionally, using an online study design ensured access to a broader range of participants, adding to the diversity of sample characteristics. While all participants were players, the type of players who participated across all five studies differed considerably. For instance, most participants from Study two were thought to be predominantly World of Warcraft players, while participants from Study five were thought to be predominantly League of Legends players. These differences in sample characteristics provide an additional component to data interpretability, wherein conclusions drawn from the study findings can be made considering the different gaming experience each user had prior to participating in the study.

All five research studies were also cross-sectional, such that data measurement took place at only one point in time, with no repeat involvement from any participant. As the aims of the thesis were to establish relationships between individual differences and design feature preference, and validating any identified relationships in subsequent studies, the repeat use of a sample was believed to detract from the generalisability of findings. Additionally, to accommodate broader analytical objectives (such as model development using an EFA and CFA), it was necessary to use a different sample for each study.

3.4 Data collection methods

3.4.1 Surveys

Another set of design similarities across all studies was the use of online surveys. In each study, participants were required to complete a range of scales, attributed to the associated cost-effectiveness, and due to their being no other viable solutions to measuring individual differences or design feature preference. For example, in-person measurement was not possible when considering the range of territories participants responded from (including users from the Europe, North America, Asia, and Africa. Additionally, some of the studies conducted took place during the Covid-19 social lockdowns of 2020-2021 which restricted face-to-face contact across most university institutions. (Study three, Study four, and Study five).

3.4.2 Demographics

Across all studies, participants completed a demographics scale, which collected general gamer characteristics, such as their country of domicile, age, gender, their favourite video games, and their years of gaming experience.

3.4.3 Design feature preference vignettes

In Study one, participants were asked to rate how representative a collection of design feature vignettes were of a given design features' primary function. In all subsequent studies, participants were asked to provide preference ratings in response to the "representative" collection of design feature vignettes (formalised later as the DFPS).

3.4.4 Individual differences

Scales which measured individual differences were also administered, with user motivation and user personality measures administered in Study two, and a Gamefulness measure administered in Study three. Given the online, cross-sectional approach, the use of surveys was appropriate in measuring these variables of interest. In addition to cost-effectiveness and limitations associated with in-person participant measures, the psychometric properties of measuring individual differences using a research scale is supported by extensive literature (McCrae & Mõttus, 2019).

3.4.5 Engagement metrics

For Study one, Study two, and Study three, user engagement was measured as part of the general gamer demographics survey, with participants asked to specify how often, and for how long they play video games on average. In Study four, user engagement was measured using a measure of task-performance, which was calculated by measuring how long the participant spent playing a series of online browser games (accessed via a web browser). In Study five, user engagement was measured via an in-game behavioural metrics scale that had been adapted to measure game-specific activity for League of Legends players.

3.5 Sample selection

Across all studies, participants were eligible to participate on the basis that they were at least 18 years of age, and that they consider themselves video games. As all participants would be asked to report their design feature preference, a degree of familiarity with gamer terms was required, as well some foundational gaming experience, so that design features they were to provide preference ratings for could be visualised and contextualised. Using the online survey design, all participants across all studies were recruited from gaming oriented Reddits. Additionally, in the first four studies, as part of the participation incentive, participants were entered into a free game giveaway, wherein the successful winner would be awarded a game of their choice, limited to $77 or £55 and purchasable only from a reputable online seller (such as STEAM, Origin, or the XBOX marketplace).

3.6 Analysis techniques

Data analysis techniques were provisionally determined by the type and extent of data collected, as well as the overarching study objectives. All analyses were statistical in nature.

3.6.1 Study one (ANOVA)

The aim of Study one was to identify which of three vignette variants was most representative of a given design features' primary function. As there were 37 design features for which there existed a separate three vignette variants, a total of 111 vignette ratings of representativeness were collected. To determine which of these variants were

most representative, a series of repeated measures ANOVAs were performed, which would determine whether there existed significant differences in representativeness between the three variants across each of the 37 comparisons. Post-hoc pairwise comparisons were then conducted to determine where these significant differences in representativeness had occurred, thereby highlighting which were vignette variants were most representative.

3.6.2 Study two (Exploratory Factor Analysis and Multiple Regression)

To fulfil the aims of Study two, two different statistical techniques were used. For the aim of furthering the model development of the design feature preference vignettes, an EFA was conducted which is a statistical technique used to reduce a high number of variables into a fewer number of factors (Hashmi et al., 2020). For the aim of identifying how design feature preference was associated with the individual differences of user motivation and user personality, a series of multiple regressions were conducted. These multiple regressions were first conducted between user engagement and design feature preference (in the form of the nine-factor model generated by the EFA), to first identify which design feature dimensions, for which preference varied, was associated with significant differences in overall user engagement. Thereafter, multiple regressions were conducted between user motivation and design feature preference, as well as between user personality and design feature preference. The aim of a multiple regression is to determine to what extent changes in a dependent variable can be attributed to changes in the independent variable. In the context of Study two, this statistical technique would identify how variance in user motivation and user personality would result in variance in design feature preference, thereby revealing which design features a user is likely to prefer based on their motivations and personality.

3.6.3 Study three (Confirmatory Factor Analysis and Multiple Regression)

To fulfil the aims of Study three, there again were two different statistical techniques used. For the aim of further developing the DFPS, a CFA was conducted, which is a statistical technique that verifies the factor structure of a set of variables (Brown & Moore, 2012). The EFA factor structure generated in Study two was therefore subject to a CFA. The suitability of performing a CFA in Study three is justified by four points. First, a CFA less suited at the early scale development stage, as it does not indicate how well

individual items (i.e., design feature vignettes) load onto each factor (Kelloway, 1995). Generally, the use of CFA in an exploratory capacity is inappropriate (Brown, 2003). Second, CFA is most suited for a new sample that was not used to generate the factor structure it seeks to verify (Fokkema & Greiff, 2017; Willmer et al., 2019). Third, a CFA is suitable to conduct on a sample of at least $n > 200$, which was far exceed in Study three (a sample size of 1111 participants). Fourth, the existing model generated in Study two scored above .3 for all design features, indicating that all variables possessed sufficient common variation to retain in the factor solution (Högberg et al., 2019).

For the aim of identifying how design feature preference was associated with the individual differences of Gamefulness, a series of multiple regressions were conducted between Gamefulness and the new DFPS model, verified by the CFA. In the context of Study three, this statistical technique would identify how variance in Gamefulness would result in variance in design feature preference, thereby revealing which design features a user is likely to prefer based on their Gamefulness value.

3.6.4 Study four (Multiple Regression)

The aim of Study four was to validate the relationship between design feature preference and user engagement, found in Study two. Whereas this initial relationship was determined when user engagement was operationalised as how long and how often they play on average, Study four was to operationalise user engagement in the form of task-performance. Participants were asked to play a series of online browser games, during which the time they spent playing would be measured. The time they had spent playing (i.e., engagement) would then be compared with their design feature preference ratings and compared with the previous relationships identified in Study two.

To determine whether the relationship between design feature preference and user engagement would persist when using task-performance measures, a series of multiple regressions were conducted between design feature preference and user engagement. In the context of Study four, this statistical technique would identify how variance in design feature preference would result in variance in user engagement, thereby revealing which design features, if preferred, will likely increase, or decrease user engagement.

3.6.5 Study five (Pearson Correlation)

The aim of Study five was to further validate the relationship between design feature preference and user engagement. The original relationship between design feature preference and user engagement was identified first in Study two, when user engagement was operationalised as for how long and how often a user plays on average. In Study four, user engagement was operationalised in the form of task-performance (Study four). In Study five however, the relationship between design feature preference and user engagement, would see user engagement operationalised as in-game behavioural metrics, which were argued as being a more objective and natural representation of player activity.

To determine whether the relationship between design feature preference and user engagement would persist when using in-game behavioural metrics, A series of Pearson's correlation tests were conducted between design feature preference and user responses to the in-game behavioural metric scale, to identify how design feature preference would relate to variances in over user engagement (operationalised as in-game behavioural metrics). The aim of a Pearson's correlation test is to determine whether there exists a linear correlation between two variables, such as whether a change in the first variable will result in a change in the second variable in the same direction (Bishara & Hittner, 2012). In the context of Study two, this statistical technique would identify whether increases in design feature preference would be associated with increases with the in-game behavioural metrics, thereby revealing whether the preference a user reported for a given design feature would be evident in their interaction with the in-game metrics of the same nature. For instance, the level of preference for a Points design feature, and the number of in-game points a user had earned.

4. Study One

4.1 Abstract

Adaptive Gamification builds on the notion that a one-size-fits-all approach to Gamification design is ineffective, given that users differ in their needs. Understanding individual differences in user needs requires an understanding of what design features are preferred by users. However, existing methods which seek to gauge user design feature preference are limited, specifically in how effectively they isolate a design feature, the terminological consistency of how the design feature is discussed, the specific detail on how a design feature takes form or is presented, the level of contextual suitability of the design feature, and the overall number of design features used in existing research. Study one sought to improve on these limitations by developing the DFPS. Following consultation of existing Gamification literature, a total of 37 design features were operationalised using vignettes. Each design feature was represented by three vignettes, all of which differed in syntactic structure and design feature context. Using an online cross-sectional design, a total of 311 players were asked to rate how representative each vignette variant was of the design feature it depicted. Using multiple t-tests to compare mean ratings of representativeness, results indicated that vignettes which were worded in the possessive form and included more design feature context achieved a significantly higher rating of representativeness than those worded in a non-possessive form and did not include design feature context. The application of the DFPS in future research is discussed, such as how it enables researchers to examine how variance in individual differences may influence or predict variances in design feature preference.

4.2 Introduction

Though there are several beneficial outcomes of Gamification, such as increased engagement within a gamified tasks or non-game contexts (Muntean, 2011), there are also instances wherein Gamification does not work (Hakulinen et al., 2013). According to Lopez and Turner (2019), a majority of cases where Gamification does not produce desired outcomes are a product of using a one-size-fits-all approach, which builds on the presupposition that the target sample of Gamification can be treated as a homogenous group that will respond alike when interacting with design features (Klock et al., 2015; Jia et al., 2016).

In contrast to this approach, a growing interest in Gamification research, referred to "Adaptive Gamification" has focused on the role of individual differences in determining the success of Gamification, recognising that any target sample of Gamification is likely to comprise of users that demonstrate variances in their preference to design feature. On this basis, the Adaptive Gamification paradigm builds on the notion that increased preference would correspond to an increase in overall user engagement with a gamified task and non-game context. For example, Codish and Ravid (2014) report that a Badge design feature increased user-perceived playfulness to a greater extent for highly agreeable personality types in comparison to low agreeable personality types.

To understand precisely how the design of Gamification could be tailored to maximise user engagement (by "targeting" relative individual differences), design features must first be operationalised to provide researchers a necessary point of reference. Primarily, this prerequisite would provide researchers with a start point (individual difference) and end point (preference for design feature), between which any relationship could be identified, scrutinised, and tested. For example, if research was interested in whether Gamification can be adaptively designed for users who score higher in trait conscientiousness, a personality scale that measures levels of trait conscientiousness would be the start point, while the end point would be the user's preference for a design feature. At present, there exists many examples of operationalised start points, with several scales that measure a broad-spectrum of individual differences, such as personality (Ashton & Lee, 2004), motivation (Kahn et al., 2015), and Gamefulness (Högberg et al., 2019). However, there are no existing operationalised measures that provide researchers with the (design feature) end point.

In the context of Adaptive Gamification, the absence of an operationalised measure for design features not only stunts further assessment of how individual differences can inform Gamification design but may also explain why the paradigm of Adaptive Gamification has only recently emerged (Böckle et al., 2017). A range of operational and conceptual barriers to operationalisation can partly explain why measures do not currently exist. Namely, these are issues concerning design feature isolation, terminological consistency, research specificity, contextual suitability, and limited scope.

4.2.1 Design feature isolation

Design features are often presented in complex combinations and are rarely isolated. For example, when presented to a user, a Leaderboard design feature, which serves the function to situate a user's progress against the progress of other users, is often presented together with a Points design feature and Rank design feature, both of which provide indicators of the user's progress (and comprises the information within the Leaderboard). In the context of Adaptive Gamification, if researchers aim to identify the extent of preference a user may possess toward a given design feature, then the first challenge is to adequately isolate the given design feature from the accompaniment of other design features. In doing so, researchers would be provided with a clear (and distinguishable) end point against which user preference and the predictive value of an individual difference, can be assessed. Failure to isolate a design feature within research could reduce the validity of any insight that may be derived from analyses which aim to identify how individual differences can influence user preference. In a research capacity, if an individual difference was found to influence a user's preference of a Leaderboard design feature, but the Leaderboard was presented with the Points and Rank design features as well, then it raises the question of how to determine which design feature (out of three) the user is reporting their preference for. Indirectly, a lack of clarity would also undermine the predictive value of any individual difference, given that it would be unclear which of the three design features (in this example) the influence of an individual difference relates to.

4.2.2 Terminological inconsistency

As a research area, Gamification is a point of convergence for several different disciplines, including Game studies, Psychology, Business and Economics, Human Computer Interaction, and Education (Deterding et al., 2011). One outcome of this are the numerous differences in classifications, definitions, and conceptual understandings that lead to inconsistent terminology and nomenclature. One case in point is the overlapping and interchangeable use of terms such Game Design Elements, Motivational affordances, and Game Mechanics. For example, the Points design feature is reported as a game design element (Barata et al., 2017), a motivational affordance (Hamari et al., 2014) and a game mechanic (Codish & Ravid, 2014; Hsu, 2017; Orji et al., 2014) respectively. Similarly, the way in which design features are conceptualised also vary, with some researchers focusing on a features' rudimentary function, or a features' intended purpose when implemented into a non-game context. For instance, Sailer et al. (2017) describe Points as being an important provider of feedback (rudimentary function), whereas Orji et al. (2014) define the Points feature as a measurement of success for in-game actions.

These inconsistencies reduce how prescriptive some research findings are. For instance, in the previous example, the conceptualisation of a Points design feature as a provider of feedback could create the potential for researchers to interpret any derivable prescription as being related to the function of serving feedback, as opposed to the Points design feature per se (the numerical representation of progress/performance). As there are alternatives design features which also provide feedback, such as a Progress Bar, the exclusiveness of any finding being related to the Points design feature could be called into question, as one could make the case that the function of feedback is of greater interest or importance.

4.2.3 Specificity

Issues of research prescriptiveness might also be considered in view of the absence of specificity, which can be defined as the degree of detail provided by researchers on design features used in research. Some of the more broadly defined design features for example, such as rewards (Jia et al., 2016), gifting, easter eggs, quests, customisation (Marczewski, 2015), personalisation, creativity tools, and social status (Tondello et al., 2017) lack important detail when reported, despite the varying way in which these features could

manifest in a non-game context. For example, the reward a user may receive in a health context could be a virtual trophy (Kukkonen & Harjumaa, 2009) or score stars (Reynolds et al., 2013), but in an occupational context a reward could be something redeemable in real life (Levy & Glimcher, 2012), such as cafeteria coupons. Moreover, in absence of detail, it is unclear whether a design feature is implemented singularly, or if a design feature comprises of more than one design feature. For example, the Creativity tools design feature is not a single design feature, rather a term used to describe a selection of singular design features, such as editing, designing, customising, and so forth.

4.2.4 Contextual suitability

More focused instances of issues relating to a lack of specificity within Gamification literature concerns the often-insufficient level of detail provided by researchers on the contextual suitability of a given design feature. Though existing use cases of Gamification predominantly use electronic applications or technologies (Johnson et al., 2016), very few distinctions are made between which design features are more effectively implemented in digital contexts, such as in a mobile application, or non-digital contexts such as a classroom or gymnasium. By design, some design features could be easier to implement in different contexts. For instance, the Leaderboard, Points, and Badge design features are easily implemented in non-digital contexts, whereas more complex design features such as Anarchic Gameplay (a user being free to do whatever they want within the game space) (Tondello et al., 2017) are likely to be much more difficult to implement.

4.2.5 Limited scope

While not a barrier to operationalisation, a broader outcome of these issues is a reduced degree of diversity in the design features that are used within Gamification research. According to Nacke and Deterding (2017), most Gamification research focuses on the utility of three main design features; Points, Badges and Leaderboards. The implementation of these design features is relatively easier than the implementation of more complex design features, potentially explaining their more frequent use. For example, Sailer et al. (2017) argue that these design features are directly visible to players, which could provide a more observable end point against which any changes in user behaviour (following interaction) can be attributed to. The greater use case for these design features might also be explained by the ease of experimental manipulation they

offer (Sailer et al., 2017) and a more generalisable level of implementation to both digital and non-digital contexts. In comparison, more complex design features, such as Behavioural Momentum (the relative gradual increase of difficulty based on the users performance level) and Item Degradation (the planned expiration of an acquired and usable item) are likely to be less generalisable to digital/non-digital contexts. As such, the scope of design feature research and the subsequent applicability of Gamification research is limited, given that a larger portion of design features are yet to be used or examined in relation to user preference (and individual differences).

A case can also be made that Gamification research must evolve and diversify the number of design features researched if it is to remain a potentially effective intervention considering the increasing scale and presence of digital technology. According to Högberget al. (2019), the existing and next generation of technology users are likely to be more familiar with video game design features. Referred to as digital natives (Granic et al., 2014; Prensky 2012; Vesa et al., 2017), these users could potentially be less effected by Gamification due to increased desensitisation and a reduction in novelty effects, as research has demonstrated low familiarity with features or schemata of a game can have a significant effect on user engagement (Li et al., 2014). As such, the diversification of design features used within research could provide a solution to the potential problem of increased user familiarity and meet the evolving characteristics of a given user base.

4.3 Rationale

The prospects of Gamification are well reported, with the existing body of literature suggesting it is an effective way to increase productivity and performance across a variety of contexts. Despite these positive findings, however, several ineffective iterations of Gamification are often reported which are understood primarily as a product of a reliance on a one-size-fits-all approach, a model that presupposes all users will uniformly respond to the Gamification of a non-game context (Lopez & Tucker, 2019). The existing reliance on the one-size-fits-all approach can be explained when considering the operational and conceptual obstacles that researchers face. Recognising this limitation, the growing area of Adaptive Gamification offers an alternative paradigm that reorients research emphasis to the role of user preference in determining user engagement, as well as the psychological antecedents that could influence user preference. To research these

aspects of Adaptive Gamification, researchers would first need to be equipped with a methodology that is designed to measure user preference for design features and also addresses broader issues of design feature isolation, terminological consistency, specificity, and overall design feature scope.

4.4 Research aim

The aim of the present study was to develop an operationalised instrument that could measure user preference for design features, which would later serve as a point of reference against which the influence of individual differences on preference could be assessed. Considering the role of design feature isolation and the broader literature issues with terminological inconsistency and specificity, the developed instrument would also need to possess high content validity and representativeness.

4.5 Methods

4.5.1 Operationalisation strategy: design feature vignettes

An effective way to operationalise design features for research is via the medium of vignettes, with which the function of a given design feature could be textually presented. Defined as "stimuli that selectively portray elements of reality to which participants respond" (Given, 2008), vignettes provide researchers with a high degree of experimental control and allow for easy design, distribution, and editing. In the context of Adaptive Gamification, operationalised design features in the form of vignettes could be presented to participants, in response to which participants could report their preference.

A benefit of vignettes is that they are relatively easy to interact with (Knauper et al., 1997), which is particularly useful in the context of understanding the relationship between individual differences and user design feature preference. To illustrate, research which focuses on individual differences, such as personality or motivation, often employ lengthy surveys which aim to capture a broad personality or motivational profile of the participant. The psychometric disadvantage of this is the duration taken to complete the survey, as well as the cognitive tax required to contemplate and reflect in response to the questions asked. Together, these issues may impair the quality of data collected (Bogen, 1996; Galesic & Bosnjak, 2009). If in the context of Adaptive Gamification research,

vignettes were to accompany measures of individual differences, it could be argued that due to their simplicity, any further cognitive load on the participant could be lessened.

Vignettes are also a cost-effective solution to isolating design features, which is a key consideration already discussed. To isolate a design feature via the medium of a vignette, a researcher would only need to adjust the textual content presented. The use of vignettes are also likely to assist in increasing the scope of design features used within research, given the high level of experimental manipulation provided and low resource requirements. By comparison, other potential methods of isolating design features could be more resource intensive or provide less experimental control. For instance, whereas vignettes present a design feature textually, an alternative approach could be visual presentation, such as in an interactive game (where a participant can directly interact with a given design feature) or in a viewable video clip (where a participant can view the design feature being interacted with). In the case of an interactive game, it is likely that researchers would need to have a purpose-built game that would enable participants to interact with isolated design features. The production of a small mobile game requires a development team, testing, software licensing, and significant finance in place (up to $150,000) (Starloop Studios, 2022). These costs are also likely to continuously increase in tandem with changing research needs, such as the design feature being presented or any development issues.

Similarly, in the case of a viewable video clip, researchers may encounter issues in design feature isolation and experimental control. To illustrate, if a video clip from an existing game was to demonstrate the function of a Progress Bar, it might be difficult to isolate that design feature from the other design features or user activity that might also be presented in the video clip. For example, a video clip demonstrating the Progress Bar being filled as a user completes in-game tasks will present not only the Progress Bar, but also the in-game tasks being completed. If such in-game tasks comprised of other design features, then researchers would be presenting more than one design feature, thus failing to achieve design feature isolation.

Being limited to conveying information textually, a key consideration in operationalising design features via vignettes would be to also focus on describing the function each design feature serves, which could ensure terminological consistency and specificity. In existing literature, researchers often omit the details of what function a design feature

serves, which prevents a more nuanced understanding of why preference is likely to vary between users, given that detail which may explain this variance is not provided. For example, the design feature of "Creativity Tools" (Tondello et al., 2017) could refer to the function of being able to customise character appearance, or alternatively to customise a profile record. While both do come under the broader term of a creativity tool, understanding whether one form is more likely to be preferred than the other is difficult to establish if such differences are not defined.

4.5.2 Development of design feature vignettes

During development of the design feature vignettes, there were three primary elements that were considered. The first, was to consider the readability of each vignette; the second was to maximise the number of design features operationalised; and the third was to ensure that the wording of each vignette focused on the function each design feature served.

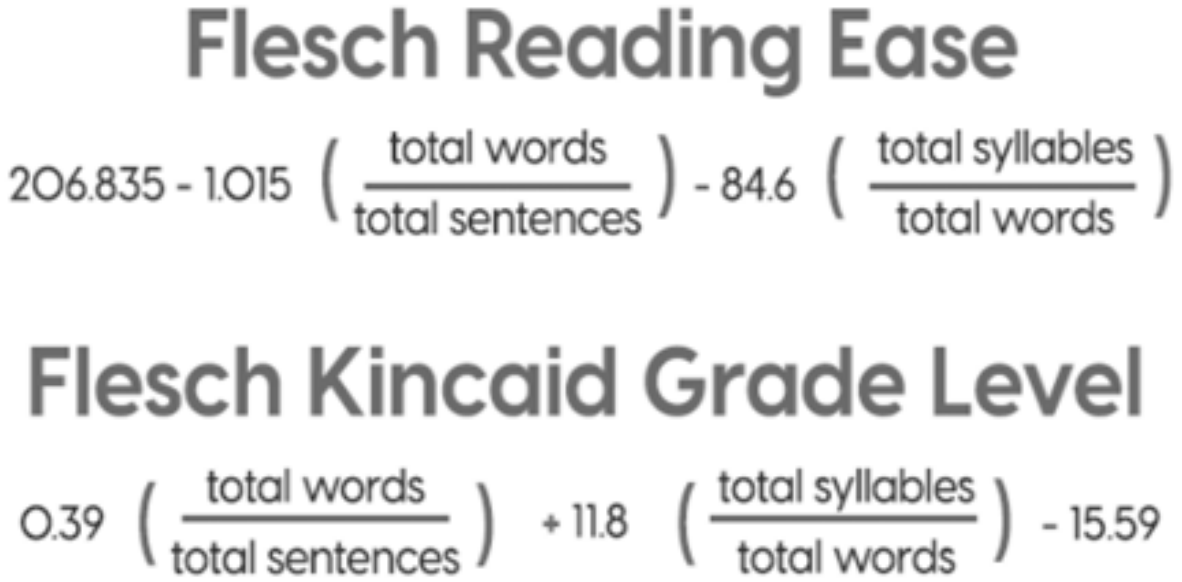
4.5.2.1 Readability

To ensure that readers would not face difficulty in reading and understanding the description of a given design feature, each vignette was considered for its level of readability. For instance, all vignettes were to be limited to 17 words, which followed a similar protocol employed by Clifford et al. (2015) who developed a standardised set of vignettes relating to situational morality. Limiting the number of words would likely reduce the number of characters each vignette comprised of, thereby increasing readability. Support for this approach comes from the Flesch Reading-Ease indices (Flesch, 1979) and the Flesch-Kincaid Grade level tests (Kincaid et al., 1975), which are tests that assign a readability score to a text depending on the number of characters and words used.

A higher score on the reading-ease test would suggest easier material, while lower scores suggest more difficult material. Inversely, a higher score on the grade level test would suggest a material is more difficult to read, whereas a lower score would suggest a material to be easier to read. The formula (Figure 4) calculates a score for each test by comparing the total number of words, sentences, and syllables with existing averages. For example, sentences with more words are more likely to receive a lower readability score, or inversely a higher-grade level. To illustrate, research has examined the readability of

newspapers based on topic, with findings suggesting areas such as Sports, Weather, and Fashion tend to score higher in indices of readability in comparison to more technical areas such as Business, Science, and Politics (Flaounus et al., 2012).

Figure 4 – Flesch reading ease and grade level formula



4.5.2.2 Number of design features operationalised

The second consideration was to increase the number of design features used within Gamification research, which builds on the previous discussions of limited scope. A total of 37 design features were identified as being suitable for operationalisation after consulting existing literature (Arnab et al., 2015; da Rocha Seixa et al., 2016; Hamari & Lehdonvirta, 2010; Lameras, 2017; Marczewski, 2015; Nacke, 2018; Orji et al., 2018; Rocha et al., 2008; Sailer et al., 2017; Tondello et al., 2017; Werbach & Hunter, 2012) and reflecting on the researcher's gaming experience and knowledge (Table 2). The criteria for inclusion were whether the function of each design feature could be isolated from the function of other design features, the design feature could be conveyed via text and did not require any visual supplement, and if the design feature could serve an implementable function in both digital and non-digital contexts. The design features which did not meet these criteria were excluded. For instance, if the design feature could not be isolated, then the research objective of understanding how preference for a given design feature is associated with variance in user engagement, would not be achievable.

Moreover, the research objective of attributing the variance in design feature preference to variances in individual differences would also not be achievable, given that any associations between preference and individual difference could not be attributed to a single design feature, as discussed in section 4.2.1.

Additionally, if design features which do not possess an implementable function were to be included in the model, then it would be difficult to attribute variance in user engagement following interaction to the given design feature. For instance, Arnab et al. (2015) lists several design features which do not possess an implementable function, and instead describe the subsequent user experience a user encounters following interaction. Some examples can be found in Table 3. As these design features do not possess implementable functions, it is unclear how any associations observed in research between design feature preference and user engagement can be replicated, given the "loose" definitions of the design features provided. An additional example can be found in research conducted by Hall et al. (2013), who propose Social Interaction as a design feature, which we argue is not an implementable design feature, but instead an experience that is curated following the implementation of different socially facilitative design features, such as, for example, a voice chat design feature, or a chat box design feature.

Table 2

List of design features found in literature which met the inclusion criteria

Design feature	**Function**	**Literature**
PVP	The element of playing against other real-life players	
Leaderboard (Competition)	Situating player progress amid the progress of other players	Sailer et al. (2017)
Complementarity	Necessary presence of other characters when completing task (for example to complete the objective the abilities of two players must be combined).	Rocha et al. (2008)
Shared Goal	Non-exclusive goals that can be accomplished more efficiently with other players	Rocha et al. (2008)
Trade	Transactions with other players in which advantages can be gained/shared (for example trading a rare item for a large sum of money)	
Text Chat	Communicating through a text channel	
Voice Chat	Communicating through a voice channel	

Emotes	Communicating through avatar behaviour with the use of emotes (for example an avatar jumping for joy)	
Trophy	Evidence of merit/achievement and indication of competency	Arnab et al. (2015)
Badges	Evidence of merit/achievement and indication of competency	Sailer et al. (2017)
Medals	Evidence of merit/achievement and indication of competency	da Rocha Seixas et al. (2016)
Tokens	Evidence of merit/achievement and indication of competency	Sailer et al. (2017)
Items	Functional objects that can be utilised to enhance skill level	da Rocha Seixas et al. (2016)
Depletion	Sanction received following failure of some sort that reduces current inventory	
Restriction	Sanction received following failure of some sort that restricts access to game features/game area	
Demotion	Sanction received following failure of some sort that demotes status	
Points	Numerical based indication of what the player has accrued	Sailer et al., (2017)
Progress Bar	Visual indication of what work is remaining before a milestone is reached/task is completed	
Leaderboard (Feedback)	Situating player progress amid the progress of other players	Sailer et al., (2017)
Walkthrough	Step by step guide on any matter that will help progress through the game (for example how to perform tasks, or the rules of the game)	Arnab et al. (2015)
Tips / Hints	Less instructional than a walkthrough, but still provides advantageous small pieces of information that can assist with the task at hand	
Notification / Prompts	Reminder or notification of changes	Lameras (2017)
Cut Scenes	Video sequences that convey story progression	Arnab et al. (2015)
Storyline	Context within which the game or characters are situated in	Arnab et al. (2015); Sailer et al. (2017)
Currency	Accumulated spendable income	Arnab et al, (2015); da Rocha Seixas et al. (2016);
Item Degradation	The planned expiration of items possessed or purchased	Hamari & Lehdonvirta (2010)
Dashboard	Platform where game history can be accessed (e.g., resources, points, achievements etc)	Lameras (2017)

Behavioural Momentum	The game gradually increasing in difficulty	Arnab et al. (2015)
Levels	Sections or parts of the game that is only accessible once a previous level is completed	Arnab et al. (2015); da Rocha Seixas et al. (2016); Lameras (2017)
Barriers	Exclusion from accessing aspects/areas of the game	
Game objectives	The end aim to complete when playing (for example in an FPS, the game objective may be to eliminate all players)	
Game goals	The smaller tasks and achievements to be fulfilled during play that will facilitate the game objective being completed	
Design / Editing / Customisation	Opportunity to design, edit or customise aspects of the game	Arnab et al. (2015)
Decision Making	Power to make decisions that affect the course of the game/story	
Avatar	A virtual model/sprite/signature representation of the gamer	Sailer et al. (2017)
Profile	Opportunity to convey aspects of oneself to other players	
Rank / Status	The assignment of a rank/status to convey the players level of skill and experience	da Rocha Seixas et al. (2016)

Table 3

Examples of design feature exclusions

Design feature	**Reason for exclusion**	**Literature**
Fun	This is a term to characterise a user experience. It is not a design feature that can be implemented, rather it is an outcome of interacting with design features.	Arnab et al., (2015)
Challenge	This is a term used to characterise parts of the user experience. It is not a design feature that can be implemented, rather it is an outcome of interacting with design features.	Arnab et al., (2015)
Urgent Optimism	This is a term used to describe the user's emotional or motivational state when playing. It is not a design feature but could be cultivated following interaction with design features.	Arnab et al., (2015)
Strategy / Planning	This is a term which refers to how a user will respond to in-game challenges. It is not a design feature, but might be supported by the use of other design features.	Arnab et al., (2015)
Protégé effect	This is a term which refers to a method of knowledge acquisition. It is not a design feature but could be facilitated with the use of other design features.	Arnab et al., (2015)
Pareto optimal	This is a term which refers to the distribution of wealth (or in-game wealth). It is not a Design feature, rather it refers more to a conceptual rule.	Arnab et al., (2015)
Mini-games	This is a term which described segments of a game that derive from the main game (a game within a game). It is not a design feature but could be set up with the use of other design features.	Arnab et al., (2015)

Realism	This is a term which describes a genre. It is not a design feature.	Arnab et al., (2015)
Virality	This is a term used to describe the potential for something to be rapidly shared and circulated over the internet (i.e., something to go viral). It is not a design feature.	Arnab et al., (2015)
Cascading information	This is a term used to describe the dissemination of information. It is not a design feature but could be supported with the use of other design features.	Arnab et al., (2015)
Role Play	This is a term which describes the voluntary adoption of in-game character personas. It is not a design feature, rather a way in which a user chooses to interact with and immerse in the game.	Arnab et al., (2015)
Ownership	This is a term which describes the sense of control a user feels when playing the game. It is not a design feature, but it can be facilitated by other design features.	Arnab et al., (2015)

4.5.2.3 Wording focus

The third consideration was to ensure that each vignette was worded to describe the function the design feature served. Given the subjectivity in determining the most effective functional description of a given design feature, part of the design process entailed the creation of vignette variants. This procedure followed similar protocols to those found in psychometric literature, such as creating a larger pool of items from which a smaller selection is drawn (Spada & Caselli, 2015). Three variants were created for each design feature (Appendix A) each differing by syntactic structure (Table 4). The first variant exclusively described the function of the design feature. The second and third variants included more context than the first to encourage participants to visualise themselves interacting with the design feature in its native environment. The difference between the second and third variants focused on changes in the possessive, which has shown to improve memory recall in related research (Shi et al., 2011).

Table 4

Differences in vignette syntactic structure

Example design feature	**Function**	**Variant 1 (context excluded)**	**Variant 2 (context included)**	**Variant 3 (context included)**
Demotion	Sanction received following failure of some sort that demotes status	The scenario in which a sanction is received following a failure - the sanction demotes status or rank	Having to receive a sanction following a failure - the sanction demotes status or rank	Being demoted and having your rank reduced after failing in some way

4.5.3 Research question

Design feature vignettes aim to describe the function a design feature serves, and in response, participants are expected to report their level of preference. As part of the validation process for this measure, three variants of vignette wording were created, each differing in design and expected effectiveness at capturing user preference. Therefore, the research question of the present study was to determine which of the three variants were rated as most representative by participants in describing the functions of design features.

4.5.4 Hypothesis

It was expected that there would be statistically significant differences in ratings of representativeness between all design feature vignette variants. It was also expected that variants which included context and were worded in a possessive for (variant 2 and 3) would score higher in ratings of representativeness than the variant which did not include design feature context (variant 1).

4.5.5 Design

The present study employed a within-subjects online survey methodology to determine which of three vignette variants were most representative of the function a given design feature serves. The functions of 37 design features were individually described across three vignette variants. Each variant was presented to participants, who in response provided a rating of representativeness using a visual analogue scale (0-100). The independent variable was the vignette variant, and the dependant variable was the level of perceived representativeness, measured by a visual analogue scale.

4.5.6 Participants

Players were recruited from the PC MasterRace Reddit, which is an online gamer community that revolves around discussion of PC gaming and comprises of 4.8 million subscribers. A total of 311 users participated. 86% of participants were male, 12% were female, and 2% chose not to identify. Ages ranged from 18 to 51 (Mean age = 30.82 years; SD = 5.91). 45% of participants were employed, 35% were students, 14% were actively seeking employment, and 1% were unemployed. Of the total sample, 50% were from

North America, 28% from Europe, 11% from Asia, and the remaining from South America, Oceania, or Africa.

4.5.6.1 Eligibility criteria

Participants were only eligible to participate if they met two main criteria. The first, was that participants must have regularly played video games for at least two hours per week, which in previous gaming research has been regarded as the minimum time spent playing to qualify as being a gamer (Kolo & Braun, 2004). It was expected that players possess the prerequisite experience to understand the functional representation of Design features, therefore ratings of representativeness would be more reliable than those given by non-players. The second criteria were for participants to have been aged 18 years or older. As the study exclusively employed online methodologies, there was no way for researchers to obtain parental consent for underage participants, therefore only those of the legal age to consent were able to participate.

4.5.6.2 Participation incentive

All participants were automatically enrolled into a free game giveaway, wherein the successful winner would be awarded a game of their choice, limited to $77 or £55 and purchasable only from a reputable online seller (such as STEAM, Origin, or the XBOX marketplace).

4.5.7 Materials

4.5.7.1 Design feature Vignettes

The function of 37 design features were described across 3 types of vignette variant (see Appendix A), totalling an overall number of 111 vignettes (Arnab et al., 2015; da Rocha Seixa et al., 2016; Hamari & Lehdonvirta, 2010; Lameras, 2017; Marczewski, 2015; Nacke, 2018; Orji et al., 2018; Rocha et al., 2008; Sailer et al., 2017; Tondello et al., 2017; Werbach & Hunter, 2012). For each of the 37 design features, three vignette variants were created. Variant 1 was worded with no design feature context, while variants 2 and 3 were worded with design feature context. Each variant was limited to 17 words and together all vignette variants achieved a mean reading-ease score of 48.44 and grade level score of 9.68, suggesting that the vignette material could be easily understood by 13–14-year-

olds, and best understood by college graduates. All variants also achieved an alpha of (*a* = .977) suggesting high internal consistency.

4.5.7.2 Visual Analogue scale

Along with the presentation of Design feature vignettes, participants were also presented with a visual analogue scale, which would be used to record perceptions of representativeness. Participants were asked "How representative are the vignettes of a [insert design feature name]" and asked to provide a preference rating to each of the vignette variants for the design feature being assessed. Representativeness was measured as a ratio variable from 1-100, as a true zero would suggest no representativeness, while higher scores would indicate greater representativeness.

4.5.7.3 Demographics

Participants were asked to submit standard demographic data, such as age, ethnicity, gender, education status, country of domicile, employment status and marital status. In addition, general gaming related demographics were also be collected, such as favourite game title, favourite game genre to play, and for how long they have played video games (in years).

4.5.8 Procedure

Participants accessed the survey via a URL link provided in the recruitment advertisement, after which they were presented with the study information sheet. After providing consent and completing the demographics sheet, participants were required to complete a trial task which introduced them to the format of the design feature rating task. In the trial task, participants were asked to rate the representativeness of a vignette which described a rabbit using a visual analogue scale (see Appendix B). Once completed, the participants were presented with 37 sets of design feature vignette variants in succession, in response to which ratings of representativeness were provided by an accompanying visual analogue scale. Once complete, participants were debriefed.

4.5.9 Ethics approval

Ethical approval was granted by Birmingham City University's research ethics committee under the Ethical Approval Code: 073.18.

4.6 Results

A series of 37 one-way repeated measures ANOVAS were conducted across 37 vignette variant sets to determine whether there was a statistically significant difference in ratings of representativeness between variants that were worded with or without context. Across all 37 repeated measures ANOVAs, there were no outliers in the data as assessed by inspection of boxplots, and the data was normally distributed as assessed by inspection of Q-Q plots. Descriptive statistics showing ratings of representativeness for each design feature vignette variant, as well as indications of normality (Skewness and Kurtosis), and the number of participants who'd provided responses is outlined in Table 5.

Table 5

Distribution statistics and mean ratings of representativeness across vignette variants

Vignette set	**Design feature**	**Mean (representativeness)**												
		Variant 1				**Variant 2**				**Variant 3**				
		Mean	**SD**	**Skew (SE)**	**Kurt (SE)**	**Mean**	**SD**	**Skew (SE)**	**Kurt (SE)**	**Mean**	**SD**	**Skew (SE)**	**Kurt (SE)**	***n***
1	PVP	76.93	21.19	-0.89 (0.14)	0.12 (0.28)	**81.09***	20.68	-1.36 (0.14)	1.60 (0.28)	68.99	26.93	-0.72 (0.14)	-0.42 (0.28)	311
2	Leaderboard (Competition)	71.14	24.46	-0.73 (0.14)	-0.17 (0.28)	**81.36***	18.1	-1.41 (0.14)	2.40 (0.28)	88.85	14.4	-1.62 (0.14)	2.37 (0.28)	311
3	Complementarity	72.39	24.26	-0.74 (0.14)	-0.26 (0.28)	74.67	25.33	-1.04 (0.14)	0.26 (0.28)	**75.39***	26.29	-1.16 (0.14)	0.55 (0.28)	311
4	Shared Goal	69.97	24.98	-0.76 (0.14)	-0.09 (0.28)	**81.80***	20.02	-1.49 (0.14)	2.08 (0.28)	78.09	22.59	-1.32 (0.14)	1.45 (0.28)	311
5	Trade	80.28	21.63	-1.47 (0.14)	1.97 (0.28)	**83.17***	19.69	-1.66 (0.14)	3.05 (0.28)	79.98	22.29	-1.33 (0.14)	1.33 (0.28)	311
6	Text Chat	85.83	17.6	-1.66 (0.14)	3.25 (0.28)	**88.00***	14.98	-1.84 (0.14)	4.69 (0.28)	66.27	25.86	-0.54 (0.14)	-0.50 (0.28)	311
7	Voice Chat	84.89	18.91	-1.65 (0.14)	2.92 (0.28)	**85.71***	17.27	-1.92 (0.14)	4.79 (0.28)	77.41	23.74	-1.22 (0.14)	0.89 (0.28)	311
8	Emotes	71.32	25.03	-0.88 (0.14)	0.03 (0.28)	75.04	23.56	-1.17 (0.14)	0.88 (0.28)	**76.36***	25.95	-1.23 (0.14)	0.79 (0.28)	311
9	Trophy	**77.44***	22.83	-1.11 (0.14)	0.75 (0.28)	74.17	23.58	-1.15 (0.14)	0.91 (0.28)	76.48	24.32	-1.25 (0.14)	1.02 (0.28)	311
10	Badges	**77.14***	22.64	-1.14 (0.14)	0.95 (0.28)	74.97	23.57	-1.10 (0.14)	0.69 (0.28)	75.56	25.17	-1.13 (0.14)	0.46 (0.28)	311
11	Medals	**76.15***	23.89	-1.16 (0.14)	0.79 (0.28)	75.5	22.68	-1.18 (0.14)	0.92 (0.28)	75.02	24.55	-1.21 (0.14)	0.88 (0.28)	311
12	Tokens	72.99	23.79	-0.83 (0.14)	-0.01 (0.28)	69.49	25.61	-0.86 (0.14)	-0.15 (0.28)	**77.13***	26.88	-1.31 (0.14)	0.75 (0.28)	311
13	Items	61.35	30.63	-0.57 (0.14)	-0.83 (0.28)	60.5	31.07	-0.56 (0.14)	-0.94 (0.28)	**72.86***	29.33	-1.08 (0.14)	0.01 (0.28)	311

14	Depletion	61.05	29.02	-0.40 (0.14)	-0.91 (0.28)	65.54	28.42	-0.68 (0.14)	-0.58 (0.28)	**77.06***	26.43	-1.33 (0.14)	0.88 (0.28)	311
15	Restriction	62.72	28.03	-0.57 (0.14)	-0.67 (0.28)	66.59	28.19	-0.72 (0.14)	-0.48 (0.28)	**72.31***	27.86	-0.98 (0.14)	-0.06 (0.28)	311
16	Demotion	70.41	26.01	-0.80 (0.14)	-0.18 (0.28)	70.06	25.06	-0.89 (0.14)	0.12 (0.28)	**81.46***	22.58	-1.48 (0.14)	1.63 (0.28)	311
17	Points	70.81	24.82	-0.75 (0.14)	-0.16 (0.28)	77.64	21.76	-1.12 (0.14)	0.84 (0.28)	**78.25***	22.95	-1.31 (0.14)	1.44 (0.28)	311
18	Progress Bar	74.5	22.9	-0.91 (0.14)	0.22 (0.28)	80.93	20.39	-1.35 (0.14)	1.59 (0.28)	**82.17***	20.67	-1.39 (0.14)	1.53 (0.28)	311
19	Leaderboard (Feedback)	73.51	23.06	-0.95 (0.14)	0.44 (0.28)	77.87	20.4	-1.22 (0.14)	1.38 (0.28)	**86.39***	17.6	-1.99 (0.14)	4.35 (0.28)	311
20	Walkthrough	77.17	23.44	-1.22 (0.14)	0.86 (0.28)	78.7	21.55	-1.21 (0.14)	0.99 (0.28)	**80.18***	21.81	-1.38 (0.14)	1.63 (0.28)	311
21	Tips / Hints	76.74	22.21	-1.16 (0.14)	0.81 (0.28)	79.47	19.67	-1.40 (0.14)	2.17 (0.28)	**80.38***	21.82	-1.31 (0.14)	1.27 (0.28)	311
22	Notification / Prompts	71.94	24.91	-0.82 (0.14)	-0.24 (0.28)	**77.77***	22.55	-1.32 (0.14)	1.45 (0.28)	73.88	25.3	-1.11 (0.14)	0.61 (0.28)	311
23	Cut Scenes	68.95	27.01	-0.70 (0.14)	-0.53 (0.28)	75.13	24.52	-1.11 (0.14)	0.46 (0.28)	**79.05***	25	-1.29 (0.14)	0.84 (0.28)	311
24	Storyline	69.87	25.27	-0.83 (0.14)	-0.12 (0.28)	**81.71***	20.32	-1.62 (0.14)	2.76 (0.28)	76.13	23.99	-1.14 (0.14)	0.68 (0.28)	311
25	Currency	78.58	22.49	-1.31 (0.14)	1.36 (0.28)	76.5	24.59	-1.28 (0.14)	0.97 (0.28)	**84.28***	20.92	-1.76 (0.14)	2.87 (0.28)	311
26	Item Degradation	70.89	27.4	-0.93 (0.14)	-0.04 (0.28)	67.26	27.01	-0.70 (0.14)	-0.36 (0.28)	**72.25***	27.86	-1.02 (0.14)	0.03 (0.28)	311
27	Dashboard	74.93	24.48	-1.08 (0.14)	0.59 (0.28)	78.48	21.23	-1.35 (0.14)	1.91 (0.28)	**81.59***	20.27	-1.66 (0.14)	3.05 (0.28)	311
28	Behavioural Momentum	**86.51***	17.66	-2.04 (0.14)	4.91 (0.28)	63.36	26.9	-0.56 (0.14)	-0.60 (0.28)	79.23	23.16	-1.49 (0.14)	1.95 (0.28)	311
29	Levels	**79.89***	23.11	-1.44 (0.14)	1.65 (0.28)	70.82	26.35	-0.92 (0.14)	-0.03 (0.28)	67.1	27.5	-0.65 (0.14)	-0.57 (0.28)	311
30	Barriers	**75.22***	25.64	-1.11 (0.14)	0.41 (0.28)	62.07	28	-0.48 (0.14)	-0.76 (0.28)	71.5	27.19	-0.87 (0.14)	-0.20 (0.28)	311
31	Game Goal	**74.61***	22.81	-1.07 (0.14)	0.77 (0.28)	74.14	23.69	-1.10 (0.14)	0.64 (0.28)	73.6	24.61	-1.11 (0.14)	0.58 (0.28)	311
32	Game Objective	**81.68***	21.49	-1.45 (0.14)	1.80 (0.28)	76.85	21.51	-1.05 (0.14)	0.50 (0.28)	74.96	24.93	-1.13 (0.14)	0.51 (0.28)	311
33	Design / Editing / Customisation	**81.12***	21.5	-1.45 (0.14)	1.87 (0.28)	80.53	21.61	-1.42 (0.14)	1.77 (0.28)	79.98	22.58	-1.56 (0.14)	2.08 (0.28)	311
34	Decision Making	**85.42***	19.02	-1.81 (0.14)	3.33 (0.28)	81.24	20.19	-1.52 (0.14)	2.31 (0.28)	84.09	19.55	-1.82 (0.14)	3.50 (0.28)	311
35	Avatar	77.58	22.66	-1.18 (0.14)	1.06 (0.28)	**80.58***	21.22	-1.56 (0.14)	2.47 (0.28)	71.08	26.35	-0.93 (0.14)	0.03 (0.28)	311
36	Profile	70.08	24.95	-0.93 (0.14)	0.32 (0.28)	**73.97***	22.5	-1.16 (0.14)	1.27 (0.28)	68.16	26.88	-0.85 (0.14)	-0.14 (0.28)	311
37	Rank / Status	79.73	21.3	-1.44 (0.14)	2.16 (0.28)	**80.47***	20.3	-1.62 (0.14)	3.05 (0.28)	71.61	26.28	-1.00 (0.14)	0.16 (0.28)	311

Notes: ***largest mean rating of representativeness***, SD = Standard deviation, Skew = Skewness, Kurt = Kurtosis, SE = Standard error, n = Sample size.

The assumption of sphericity was violated in most cases (32 out of 37 tests), therefore where applicable a Greenhouse-Geisser (1959) correction was applied (see Table 6). Table 6 also highlights where there existed statistically significant differences across all 37 tests. The results indicated that there was a statistically significant difference between ratings of representativeness across most variant sets (shown to be in 30 out of 37 tests), partially supporting the hypothesis that there would be statistically significant differences in ratings of representativeness between design feature vignette variants.

Table 6

Greenhouse-Geisser corrections and significant differences within vignette variant sets

Vignette variant set	**Design feature**	**Mauchly's test of sphericity**		**Greenhouse-Geisser correction (ε)**	***df***	***F***	***P value***	**Significant**
		χ^2 (2)	***p***					
1	PVP	6.40	*p < .050*	0.98	1.96, 607	34.16	***p < .001****	*Yes*
2	Leaderboard (Competition)	52.56	*p < .001*	0.87	1.73, 536	98.07	***p < .001****	*Yes*
3	Complementarity	5.18	*p > 050*		2, 620	1.65	*p > 050*	*No*
4	Shared Goal	5.91	*p > 050*		2, 620	28.85	***p < .001****	*Yes*
5	Trade	19.66	*p < .001*	0.94	1.88, 583	3.48	***p < .001****	*Yes*
6	Text Chat	45.17	*p < .001*	0.88	1.76, 545	144.77	***p < .001****	*Yes*
7	Voice Chat	76.18	*p < .001*	0.82	1.64, 508	23.7	***p < .001****	*Yes*
8	Emotes	52.34	*p < .001*	0.87	1.73, 536	5.4	***p < .050****	*Yes*
9	Trophy	31.60	*p < .001*	0.91	1.82, 565	2.34	*p > 050*	*No*
10	Badges	25.77	*p < .001*	0.93	1.85, 574	1.11	*p > 050*	*No*
11	Medals	13.43	*p < .001*	0.96	1.91, 594	0.28	*p > 050*	*No*
12	Tokens	14.36	*p < .001*	0.96	1.91, 593	10.66	***p < .001****	*Yes*
13	Items	26.95	*p < .001*	0.92	1.84, 572	36.51	***p < .001****	*Yes*
14	Depletion	60.51	*p < .001*	0.85	1.69, 526	49.33	***p < .001****	*Yes*
15	Restriction	66.32	*p < .001*	0.84	1.67, 519	20.37	***p < .001****	*Yes*
16	Demotion	67.00	*p < .001*	0.84	1.67, 518	41.85	***p < .001****	*Yes*
17	Points	20.41	*p < .001*	0.98	1.88, 582	22.92	***p < .001****	*Yes*
18	Progress Bar	30.04	*p < .001*	0.87	1.83, 567	17.89	***p < .001****	*Yes*
19	Leaderboard (Feedback)	109.06	*p < .001*		1.54, 477	61.1	***p < .001****	*Yes*
20	Walkthrough	11.31	*p < .050*		1.91, 598	2.52	*p > 050*	*No*

21	Tips / Hints	70.77	*p < .001*	0.94	1.66, 514	4.25	***p < .050****	*Yes*
22	Notification/ Prompts	43.75	*p < .001*	0.88	1.76, 547	8.31	***p < .050****	*Yes*
23	Cut Scenes	6.88	*p < .050*	0.98	1.95, 606	19.66	***p < .001****	*Yes*
24	Storyline	21.57	*p < .001*	0.97	1.87, 580	29.19	***p < .001****	*Yes*
25	Currency	1.01	*p > .050*		2, 620	14.318	***p < .001****	*Yes*
26	Item Degradation	17.33	*p < .001*	0.95	1.89, 587	5.59	***p < .050****	*Yes*
27	Dashboard	117.13	*p < .001*	0.76	1.52, 471	13.5	***p < .001****	*Yes*
28	Behavioural Momentum	4.22	*p > .050*		2, 620	125.27	***p < .001****	*Yes*
29	Levels	12.91	*p < .050*	0.96	1.92, 595	28.92	***p < .001****	*Yes*
30	Barriers	14.92	*p < .001*	0.96	1.91, 592	34.26	***p < .001****	*Yes*
31	Game Goal	6.85	*p < .050*	0.98	1.95, 606	0.26	*p > 050*	*No*
32	Game Objective	35.25	*p < .001*	0.90	1.80, 559	12.13	***p < .001****	*Yes*
33	Design / Editing / Customisation	22.28	*p < .001*	0.94	1.87, 579	0.44	*p > 050*	*No*
34	Decision Making	20.40	*p < .001*	0.94	1.88, 582	6.55	***p < .050****	*Yes*
35	Avatar	15.84	*p < .001*	0.95	1.90, 590	22.02	***p < .001****	*Yes*
36	Profile	30.73	*p < .001*	0.91	1.82, 566	10.24	***p < .001****	*Yes*
37	Rank / Status	13.96	*p < .001*	0.96	1.91, 593	24.68	***p < .001****	*Yes*

Notes: *largest mean rating of representativeness

Post-hoc pairwise comparisons summarised in Table 7 highlight the differences between vignette variants for which ratings of representativeness were significantly different. The results indicated that in most cases, ratings of representativeness were significantly higher for vignettes which included context (variant 2 and 3) e.g., *Being able to complete an objective more effectively if you work with another player*, than vignettes which did not include context (variant 1) e.g., *The scenario where the completion of an objective is more likely if working together with another player*. For example, variant 3 scored the highest mean most frequently (for 17 design features), compared to variant 1 which scored the lowest mean most frequently (for 15 design features).

Table 7

Post-hoc pairwise comparisons between vignette variants for each design feature

Vignette set	Design feature	Significant differences		
		Variant 1 and 2	Variant 1 and 3	Variant 2 and 3
1	PVP	**p < .050**	**p < .001**	**p < .001**
2	Leaderboard (Competition)	**p < .001**	**p < .001**	**p < .001**
3	Complementarity	p > .050	p > .050	p > .050
4	Shared Goal	**p < .001**	**p < .001**	**p < .050**
5	Trade	**p < .050**	p > .050	**p < .050**
6	Text Chat	p > .050	**p < .001**	**p < .001**
7	Voice Chat	p > .050	**p < .001**	**p < .001**
8	Emotes	**p < .050**	**p < .050**	p > .050
9	Trophy	**p < .050**	p > .050	p > .050
10	Badges	p > .050	p > .050	p > .050
11	Medals	p > .050	p > .050	p > .050
12	Tokens	**p < .050**	**p < .050**	**p < .001**
13	Items	p > .050	**p < .001**	**p < .001**
14	Depletion	**p < .001**	**p < .001**	**p < .001**
15	Restriction	**p < .001**	**p < .001**	**p < .001**
16	Demotion	p > .050	**p < .001**	**p < .001**
17	Points	**p < .001**	**p < .001**	p > .050
18	Progress Bar	**p < .001**	**p < .001**	p > .050
19	Leaderboard (Feedback)	**p < .001**	**p < .001**	**p < .001**
20	Walkthrough	p > .050	**p < .050**	p > .050
21	Tips / Hints	**p < .050**	**p < .050**	p > .050
22	Notification / Prompts	**p < .001**	p > .050	**p < .050**
23	Cut Scenes	**p < .001**	**p < .001**	**p < .050**
24	Storyline	**p < .001**	**p < .001**	**p < .001**
25	Currency	p > .050	**p < .001**	**p < .001**
26	Item Degradation	**p < .050**	p > .050	**p < .001**
27	Dashboard	**p < .050**	**p < .001**	**p < .001**
28	Behavioural Momentum	**p < .001**	**p < .001**	**p < .001**
29	Levels	**p < .001**	**p < .001**	**p < .050**
30	Barriers	**p < .001**	**p < .050**	**p < .001**
31	Game Goal	p > .050	p > .050	p > .050
32	Game Objective	**p < .001**	**p < .001**	p > .050
33	Design / Editing / Customisation	p > .050	p > .050	p > .050
34	Decision Making	**p < .001**	p > .050	**p < .050**
35	Avatar	**p < .050**	**p < .001**	**p < .001**
36	Profile	**p < .001**	p > .050	**p < .001**
37	Rank / Status	p > .050	**p < .001**	**p < .001**

The final selection of vignette variants for all 37 design features can be found in Table 8. Additionally, readability scores for the 37 final vignettes (assessed by the Flesch-reading ease test and the Flesch-Kincaid grade level tests) are described in Table 9. For the reading ease test, the average score was 51.26, suggesting that the vignettes would be easily understood by 13–15-year-olds and best understood by college graduates. For the grade level test, the average score of 9.23, indicating that the vignettes would be easily read by 13–14-year-olds.

Table 8

Final selection of 37 vignette variants

Vignette set	Mechanic	Vignette variant	Vignette
1	PVP	2	Being able to compete against other players
2	Leaderboard (Competition)	3	Being able to see how your score and rank compare with other players
3	Complementarity	3	Being required to complete an objective or task only with the help and abilities of another player
4	Shared Goal	2	Being able to complete an objective more effectively if you work with another player
5	Trade	2	Being able to trade inventory items/currency with other players in exchange for items/currency
6	Text Chat	2	Being able to communicate with other players via text chat
7	Voice Chat	2	Being able to communicate with other players via voice chat
8	Emotes	3	Being able to express your emotion and feelings through your avatar behaviour (such as jumping or dancing)
9	Trophy	1	The scenario in which a trophy is received after completing an achievement or milestone
10	Badges	1	The scenario in which a badge is received after completing an achievement or milestone
11	Medals	1	The scenario in which a medal is received after completing an achievement or milestone
12	Tokens	3	Being able to earn tokens after completing challenges, that can be used to buy game content
13	Items	3	Being able to receive items after completing challenges or tasks
14	Depletion	3	Having to lose points, items or currency after failing in some way
15	Restriction	3	Having to lose access to some aspects of the game after failing in some way
16	Demotion	3	Being demoted and having your rank reduced after failing in some way
17	Points	3	Being able to see your progression in a number format

18	Progress Bar	3	Being able to see how close you are to reaching a milestone in a bar format
19	Leaderboard (Feedback)	3	Being able to see how your score and rank compare with other players
20	Walkthrough	3	Having the option to receive a step-by-step guide on how to complete tasks or play the game
21	Tips / Hints	3	Being able to receive tips and hints when playing
22	Notification / Prompts	2	Being able to receive key notifications and updates when playing
23	Cut Scenes	3	Being able to see how the story progresses through cutscenes
24	Storyline	2	Having a central theme and story that the game revolves around
25	Currency	3	Being able to spend your in-game money/currency on game content
26	Item Degradation	3	Knowing that you must use some items/game content sooner rather than later due to time expiration
27	Dashboard	3	Being able to access game information, such as your game history, resources, profile, friends list, achievements etc
28	Behavioural Momentum	1	The gradual increase in difficulty as the game goes on
29	Levels	1	Sections or parts of the game that are only accessible once a previous/existing level is completed
30	Barriers	1	The exclusion from accessing specific aspects or parts of the game
31	Game Goal	1	The smaller and more immediate goals that once completed, will assist in fulfilling the game objective
32	Game Objective	1	The overarching goal when playing a game mode
33	Design / Editing / Customisation	1	The option to edit or design aspects of the game (e.g., avatar, environment, inventory)
34	Decision Making	1	The power to make decisions that significantly alter the course of the game/story
35	Avatar	2	Being able to represent yourself via a virtual model/sprite/signature
36	Profile	2	Being able to immediately convey aspects of yourself to other players
37	Rank / Status	2	Being assigned a category and rank that reflects your ability, score and/or experience

Table 9

Readability scores for 37 design feature vignettes

Readability test	**Number of Design features**	**Mean**	**Std. Deviation**
Flesch-reading ease score	37	51.26	22.10
Flesch-Kincaid grade level score	37	9.23	3.15

4.7 Discussion

The aim of the present study was to develop a research instrument that could measure user preference for design features. 37 design features were individually operationalised across three vignette variants, each of which differed by wording and degree of context included. It was hypothesised that there would be statistically significant differences between the three vignette variants for each design feature that was operationalised. The results revealed significant differences in ratings of representativeness between the three vignette variants in most cases, therefore partially supporting the hypothesis. It was also hypothesised that vignette variants which were worded in a possessive form and included more design feature context would score higher in ratings of representativeness, than variants which were not worded in a possessive form and did not include design feature context. The results supported this hypothesis, with variant 2 and 3 respectively scoring higher in ratings of representativeness than variant 1. The most representative items were then compiled into a scale that would measure user design feature preference, referred to at the DFPS.

Out of 37 comparisons, there were 27 instances where the vignette variant worded in a possessive form were rated as more representative by participants than those which were not. This could be explained by a possessive-benefactive connection (Litchenberk, 2002). The possessive-benefactive is a polysemy (co-existing meanings for a word of phrase) which suggests that the possessive form can often be closely associated with the process of becoming a beneficiary or recipient of something. In the context of this study, vignettes worded in the possessive form (e.g., being able to see how your score and rank compare with other players) were rated as more representative by participants. It could be argued that more positive appraisals of representativeness were provided by participants in response to reading possessive form vignettes, due to the possessive-benefactive connection, such that participants were primed to view the design feature as a reward or benefit, therefore provided higher ratings. In support, researchers have also found that participant perceptions on their preference for a given stimuli was strongly influenced by whether the stimuli (e.g., a cup or piece of bread) was presented in the possessive form (e.g., my cup), with possessive pronouns being more frequently associated with more positive appraisals of preference (Shi, Zhou, Han, Zhang, and Liu, 2011).

4.7.1 Study Improvement suggestions

A limitation of the present study was the number of design features that were operationalised. In total, the DFPS comprised of 37 design feature vignettes. It could be argued that this does not encapsulate all design features that can be operationalised and applied in Gamification research. For example, shortly after the present study was conducted, published research highlighted additional design features which could be included in the DFPS, such as the Punish design feature (being able to punish an opponent once you have defeated them) and the Vote design feature (having the opportunity to vote on something (e.g., map, weapon, rules) (Ferro, 2018). Given the rapidly evolving area of Gamification (Koivisto & Hamari, 2019), such advances can quickly emerge, therefore a future direction from this study would be to revise the DFPS and potentially expand the overall number of items included.

Future research might also benefit from recognising the readability scores achieved by the design feature vignettes. The present study adopted the approach that the readability of the vignettes could be defined in terms of general reading ability of age groups. Findings identified that the design feature vignettes were suitable for readers above the age of 13, which indicates that individuals with the reading ability of the average 13-year-old would not face any difficulty in understanding the content of the design feature vignette. However, readers below the age of 13, who would not be expected to possess that level of reading ability may find the design feature vignettes difficult to interpret, therefore the application the DFPS might not be suitable for younger participants. Age has been found to play a significant role in shaping the game preferences of players, as well as how receptive users are to different game design strategies or interventions (Birk et al., 2016; Blocker et al., 2014). It is also likely that design feature preference, which is the metric measured by the DFPS, might also be influenced to some extent by user age. Given the unsuitability of using the DFPS with younger users, future research could revisit the wording of the DFPS and seek to improve the readability of the vignettes, such that they are more applicable with younger users (though this is outside of the scope of the PhD).

To make the vignettes more accessible to users, researchers might also consider developing context-specific iterations of DFPS. Studies indicate that familiarity with a given stimuli or context can promote better recall and visualisation, especially in the

context of video game stimuli (Martí-Parreño et al., 2017). In the case of the DFPS, if readers are supported in their recall or visualisation of a given vignette, it could be argued that any corresponding rating of preference provided would be more representative of the users' preference, as a clearer picture of the design feature is generated from the participant's perspective. One suggestion of a context specific DFPS might be to contextualise the vignettes to a given game. As an example, a World of Warcraft DFPS would include specific references to aspects of the game. Some suggestions are provided in Table 10. Notably, the inclusion of more game specific information could increase the number of words the vignettes comprise of, thus considerations must be made on how this would impact overall user readability.

Table 10

Context specific suggestions of the DFPS: World of Warcraft

Design feature	**Original vignette**	**Context specific vignette (WoW)***
Rank / Status	Being assigned a category and rank that reflects your ability, score and/or experience	Being assigned a category and rank that reflects your ability, score and/or experience, **e.g., reaching the highest-level cap in the current expansion (currently Shadowlands) such as level 60**
Restriction	Having to lose access to some aspects of the game after failing in some way	Having to lose access to some aspects of the game after failing in some way **e.g., when you die you lose access to the game for 6 minutes**
Item Degradation	Knowing that you must use some items/game content sooner rather than later due to time expiration	Knowing that you must use some items/game content sooner rather than later due to time expiration **e.g., only being able to use the Chromatic Sword for 10 days after you acquire it**
Trade	Being able to trade inventory items/currency with other players in exchange for items/currency	Being able to trade inventory items/currency with other players in exchange for items/currency **e.g., trading spell stones, firestones, or devices with other players**

Notes: *text in the Context specific (WoW) column that is highlighted in bold is where the design feature vignette

A further improvement might also be to have included participants from more than one gaming community. The present study recruited all participants from the /r/PCMASTERRACE, which is an online gaming community that orients discussion on PC gaming, often via comparisons to the "inferior" mode of platform gaming (such as those on dedicated sytems such as XBOX, Playstation and Nintendo) (Hartup, 2015). Although the community does not focus on a specific title or genre, one could argue that most users

from the community (and therefore most participants in the present study) play PC games. Research has highlighted differences in users' perception, preference, and experience of gaming based on the platform on which they play (Omar et al., 2011), with significant differences in how the game operates, and the user experience it can provide. For example, playing games on a PC tends to offer greater graphic capabilities, when compared with that which is offered by console or mobile games. However, PC systems capable of playing games can often be expensive, and do not offer the degree of play flexibility and freedom that can be experienced with mobile games, such as not requiring a power source. Similarly, console systems often require additional equipment, such as a monitor or joystick / controller, and almost exclusively function as a games console, whereas the use case for mobile or PC platforms often extend beyond that of playing video games (Geraldus, 2015). Research has also highlighted that differences in these aspects, such as game graphics, can influence the user experience, with higher end graphics supporting a more immersive experience (McMahan, 2013). Given these differences, it could be argued that focusing on a gaming community which almost exclusively plays on one gaming platform might have skewed results, such that corresponding insights might not be as applicable to they might not be applicable to users who primarily play other platforms (such as mobile games, or console games).

4.7.2 Summary

The primary outcome of the present study was an operationalised scale (the DFPS) that could measure design feature preference and enable further investigation on how it might be related to individual differences. Extensive attention has focused on how individual differences, such as user motivation (Przybylski et al., 2010) and personality (Zammitto, 2010) could influence player behaviour and game preference. In the context of Gamification, the DFPS enables the investigation of how design feature preference relates to individual differences, such that the design of Gamification can be tailored or adapted to the individual needs of the user. Thus, the DFPS offers a novel contribution to the field of Adaptive Gamification by providing measures with which user needs can be better understood to inform Gamification design. As part of further psychometric development, the DFPS would benefit from potential expansion, such that other design features which were not measured at this time are included. Additionally, the application of the DFPS in further studies is critical in its assessment of validity and reliability.

5. Study Two

5.1 Abstract

A fundamental objective of the Adaptive Gamification approach is to tailor design to meet the varying needs of the end user, such that the user experience is more suited to the user and likely to increase engagement. Two areas of particular importance are user motivation and user personality, which existing gaming research indicates possess a strong influence on user engagement and game preference. Study two sought to understand how user motivation and user personality would predict variance in design feature preference, thereby providing guidance on how user needs can be more effectively fulfilled by design. Using an online cross-sectional design, a total of 900 players completed the DFPS, along with the Trojan Player Typology Gaming Motivation Scale (Kahn et al., 2015), and the HEXACO-60 (Ashton & Lee, 2009) personality test. The DFPS was subject to an EFA, which returned a nine-factor solution, sorting all design features into nine factors: Difficulty, Competition, Reward, Accessibility, Loss, Expression, Cooperation, Improvement, and Narrative design. Multiple regression analysis highlighted a significant association between user motivation and user personality on design feature preference; specifically, the motivation of Story-driven, Socialisers, and Escapist, and the personality traits of Extraversion, Agreeableness, and Emotionality. The implications of this study are discussed, such as how the DFPS would benefit from further model development in the form of a CFA.

5.2 Introduction

The previous study aimed to operationalise design features to enable more detailed assessment of how Gamification could be adapted to meet the needs of the user, as per their design feature preference. Central to this aim was the development of the DFPS, which is a research instrument that can measure how much a user prefers a given design feature (thus indicating which design features are best suited to maximise engagement for each user). The present study sought to further validate the DFPS model, by applying it to the field of individual differences, such that user preference could be assessed in relation to the users' personal characteristics, which has broadly been found to influence the way in which users interact with games or gamified systems (Figure 5). Two areas of individual differences which are strongly supported within gaming literature as influencing user engagement are motivation and personality.

5.2.1 Motivation

Heller (2012) suggests that variance in how users engage with video games can be attributed to why they play. Though research in this area is extensive (Bartle, 1996; Dmetrovics et al, 2011; Sherry et al., 2006; Wan & Chiou, 2007; Yee et al., 2012), there appears to be two main perspectives which emerge both of which differ by underlying epistemology. The first, is general motivation theory which suggests that video games are one (of many) contexts within which general motivations can be fulfilled (e.g., Wan & Chiou, 2007; Lafreniere et al., 2012). The second, is gaming motivation theory, which suggests that videos games provide the only context within which game-related motivations can emerge, develop, and be fulfilled (e.g., Bartle, 1996; Lee et al., 2012; Yee et al., 2012). In other words, general motivation *can* be fulfilled within a video game context, whereas gaming motivations *can only* be fulfilled within a video game context. Given these differences, understanding the characteristics of each approach could help inform how motivation can be considered when adapting Gamification to the needs of the user.

5.2.1.1 General motivation theory

General motivation theories are applied to understand human behaviour in various domains and areas within psychology, and are developed independent to the behaviour

of players or behaviour observed within game contexts. For example, self-determination theory (Deci & Ryan, 1980) has been used to explain alcohol use (Knee & Neighbors, 2002), healthy eating (De Man et al., 2020), and learning (Jang et al., 2016). Given the wide-ranging areas of application, self-determination theory can therefore be regarded as a more generalisable motivation theory applicable to several areas of human behaviour, one of which is video gaming. To illustrate, aspects of extrinsic motivation have been used to understand gamer behaviour, such as Integrated regulation (where a behaviour is performed out of necessity) that is demonstrated by some players' tendency to play games that align with career goals (such as game design). Similarly, Introjected regulation (behaviour performed to supress internal pressures) is demonstrated by individuals playing games to supress frustration or restlessness (Lafreniere et al., 2012). Wan and Chiou (2006) adopted a humanistic needs perspective and concluded that players who reported higher levels of video game engagement also reported stronger motives of belongingness and self-esteem, whereas lower levels of video game engagement corresponded to users who reported stronger motivations to self-actualise (reaching one's full potential; Collins, 2007).

Though various theoretical perspectives have been used to explain the motivations of users, given the unique features, degree of interactivity, and potential for immersion afforded by video games (Wang et al., 2009), it could be argued that a deductive approach to explain user motivation within a gaming context is unsuitable. Indeed, applying general motivation theory to interpret the motivational needs of users could prevent a nuanced understanding of gaming motivation. Take for example the Killer motive (Bartle, 1996), which is defined as the motivation to ruin the play experience of other players. One could argue that this type of motive, which is context-dependant to video games, could be misinterpreted or inaccurately defined if it was to be understood from a general motivation perspective. For instance, the motive to ruin the play experience of others could be misconstrued as a user demonstrating psychopathic traits, such as a lack of empathy or desire to inflict pain to others (White, 2014). Thus, general motivation theories lack a degree of nuance and ethnography needed to accurately represent the motives of players. In the context of understanding the relationship between user design feature preference and the individual difference of motivation, general motivation theory may be limited in the scope of insight it provides to understand the nuanced aspects of motivations present in gaming contexts.

5.2.1.2 Gaming motivation theory

In contrast to general motivation theory, gaming motivation theory adopts an inductive approach, wherein the motivations that drive user behaviour are derived from within the game context, such that they are more game-specific and less generalisable to non-game contexts (Khan et al., 2015). Bartle (1996) first introduced the concept of gaming motivation theory by taxonomising users into player types based on their in-game behaviour and play preferences (what they like to do and the features they make the most use of) which were thought to be underpinned by their motives to play. To illustrate, Bartle's (1996) player taxonomy comprised of four player types, each characterised by distinct objectives (or motivations); Achievers, Explorers, Killers, and Socialisers (see Table 11).

Other key gaming motivation theories include that posed by Sherry (2006), who proposed six dimensions of gaming motives; Competition, Challenge, Social interaction, Diversion, Fantasy, and Arousal, based on focus groups and questionnaire data from over 1300 US students. One of the most cited contributions to gaming motivation theory comes from the seminal work of Yee, Ducheneaut and Nelson (2012), who introduced the Motivations for Online Play taxonomy. Building on previous work (Yee, 2006), the updated model draws on cross-cultural data from over 3000 World of Warcraft players and proposes that player motivation can be categorised into three overarching motivations, each of which comprising of sub-component motivations (see Table 12). Yee et al. (2012) suggest that players do not exclusively foster one motivation or the other, such that one player can exhibit motivational tendencies of more than one motivation at any given time (much like the fluid nature of other individual differences, such as personality; Harris et al., 2016).

Table 11

Bartle's MUD player types (Bartle, 1996)

Socialisers	**Explorers**	**Achievers**	**Killers**
Motivated to build relationships with others	Motivated to explore the virtual world	Motivated to achieve in-game goals and accumulate	Motivation to impose and ruin the experience of others

Table 12

Yee's MMORPG gaming motivations and sub-component motivations (Yee et al., 2012)

Achievement	Social	Immersion
Advancement	**Socialising**	**Discovery**
Progress, Power, Accumulation, Status	Casual chat, Helping Others, Making friends	Exploration, Lore, Finding hidden items
Mechanics	**Relationship**	**Role-Playing**
Numbers, Optimisation, Templating, Analysis	Personal, Self-Disclosure, Find and give support	Story line, Character history, Roles, Fantasy
Competition	**Teamwork**	**Customisation**
Challenging others, Provocation, Domination	Collaboration, Groups, Group achievement	Appearances, Accessories, Style, Colour schemes
		Escapism
		Relax, Escape, and avoid real life problems

Many of these models are, however, limited in applicability, primarily because they are derived from research that focused exclusively on one game genre. For example, Bartle's (1996) taxonomy was based on players exclusively from a multi-user dungeon (MUD) genre, while the Yee et al. (2012) model was based on players exclusively from the massively multiplayer online role-playing game (MMORPG) genre. MUDs offer a real-time, textually based world, where individuals can play against others in the context of a virtual and fictional context, and within these worlds' players can compete against one another, explore their environment, and complete quests. MMORPGs can be regarded as somewhat similar but often with many more features available to the player, such as advanced graphics, story themes, and customisability (Paik & Shi, 2013). MUDs are often text-based where all interactions are performed via written commands, whereas MMORPGs are usually in third person and interactions are performed through a keyboard and mouse, or control pad control scheme. These differences are likely to lead to different user experiences when playing each genre given the presence (or absence) of some design features over others. As such, the motivations that are likely to emerge within each game space will likely be more internally unified but externally differ from other game spaces, and although each model will possess a strong capacity to explain the motivational tendencies of players of their respective genres, they are limited in their comprehensiveness to explain gaming motivations for other genres.

These differences highlight that game genres often comprise of unique and relative design features, which can lead to differences in how the gamer interacts with the game (Griffiths & Nuyens, 2017). To illustrate, Ghman and Griffiths (2012) report that game genre is likely to mediate player motivation and level of game engagement, with those who play first-person shooters being more motivated by achievement than those who play real time strategy, and that those who play genres which offer extensive role-play opportunities reporting greater engagement than those who play first-person shooters or real time strategy genres. As such, one can expect variance in player motivation and the degree of engagement they direct toward a video game as being influenced in some capacity by the game they are playing (Dieris-Hirche et al., 2020).

On the basis that a given game genre will facilitate players to foster some gaming motivations over others, and that the variance in motivations fostered between players of different genres can also determine their level of game engagement, understanding how genres differ becomes increasingly important. One suggestion is the design features a given genre comprises of. For example, studies suggest that that Massively Multiplayer Online Role-Playing Games (MMOPRGs) are considered most addictive genre of games due to the design features they emphasis (Scott and Porter-Armstrong, 2013). To illustrate, most MMORPGs strongly encourage social interaction with other players (i.e., being able to play in guilds, and having to complete in-game objectives with other players) which are design features understood to significantly increases user retention and engagement (Kuss, 2013). Similarly, MMORPGs also facilitate an extensive level of customisation or role-play, which are design features closely associated with the achievement of user immersion and increased user engagement (Bowman, 2018; Ducheneaut et al., 2007). Taken into consideration, if each genre differs in-part by the design features with which it is comprised, then the logical extension of this would be to regard motivation findings from studies that focus only on one game genre sample as being less applicable than those which focus on more than one game genre group. To support, Rigby (2004) argues that Bartle's taxonomy only reflects the MUD genre's structure, dynamics and content, while Kahn et al. (2015) states that MMORPGs do not represent all games. This position closely coincides with the fundamental position of the Adaptive Gamification approach; the recognition that not all players are the same (Yee, 2006). As such, in the context of understanding the relationship between user design preference and the individual difference of motivation, the usage of models or

frameworks that do not focus exclusively on one genre is a more appropriate application of gaming motivation theory.

One such example of this is the Trojan Player Typology (Kahn et al., 2015), which is a gaming motivation model based on two popular game genres: Massively multiplayer online games (MMO), and Multiplayer online battle arenas (MOBA). MMOs are a genre of games which facilitate many players to play simultaneously much like a MMORPGs. However, games of this genre are not exclusively focused on providing a role-play experience. MOBAs are a strategy game in which two opposing teams compete against each other in a predefined battlefield/arena. Both genres are widely regarded as two of the most popular genres played. For example, the active monthly player base for the MOBA, *League of Legends*, was 80 million monthly players (27 million daily players), while the MOBA *PlayerUnknown Battlegrounds* (known by the acronym *PUBG*) reported the second highest player account of all time, at 804 million players (Chapple, 2020). The model proposes six dimensions of gaming motivation described through a player typology, with each category defining the primary motives for play (see Table 13).

Unlike other gaming motivation models, the Trojan typology benefits from being derived from more than one game genre (MMOs and MOBAs), thereby increasing its generalisability to more than one gamer population. Moreover, the model was cross-culturally validated with eastern participants from China and western participants from North America and was developed using data from over 41000 players which potentially improved the extent of coverage for variance in player motivation. The model was also behaviourally validated, such that self-reported motivational tendencies of players were cross-referenced with their in-game behaviour. For example, players who reported strong socialiser motives would also be assessed by how many players they would play with (average team size), with expectations that stronger socialiser tendency would correspond to larger team size than those with weaker socialiser tendencies.

Table 13

Kahn et al. (2015) Trojan Player Typology

Socialisers	**Completionists**	**Competitors**	**Escapists**	**Story-driven**	**Smarty-pants**
Play to socialise	Play to complete every aspect of the game	Play to succeed	Play to escape from real life	Interested in the game story/narrative	Play for intellectual stimulation and challenge

5.2.2 Personality

In the case of personality, most work has approached the relationship between user personality and game engagement from a big five perspective (Graham and Gosling, 2013), which proposes five dominant personality trait dimensions; neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness. Several studies report variance in game engagement tends to correlate with higher or lower scores in some of these personality dimensions.

For example, Neuroticism, which is the trait concerned with tendencies towards negative feelings (e.g., anxiety and self-doubt) is often found to be associated with higher levels of user engagement (Lehenbauer-Baum et al., 2015; Walther et al., 2012). One interpretation of this is that those who are more inclined to experience negative feelings may pursue video game experiences as a form of real-life escapism (de Hesselle et al., 2021).

Levels of Extraversion, which is the trait concerned with tendencies for social interaction, is also understood to also predict variance user engagement, with those who are lower in Extraversion tending to demonstrate higher levels of user engagement (de Hesselle et al., 2021; Müller et al., 2014). Similarly, lower levels of Agreeableness (the trait concerned with prosociality and social harmony), Conscientiousness (the trait concerned with individual diligence), and Openness (the trait concerned with open-mindedness and creativity), all predict higher levels of user engagement (de Hesselle et al., 2021; Wang et al., 2014). In the context of Gamification, the relevance of these findings and how each trait may predict variance in user engagement relates to whether the trait tendencies of users can be leveraged to adapt and curate a more engaging experience.

5.2.2.1 HEXACO model of personality

Though the big five perspective has provided a stable and widely used measure for user personality, a criticism often wagered is that the model is restrictive in not including other nuanced personality dimensions (Boyle, 2008). In the context of gaming and understanding user design preference, other models may prove more effective in application (Zeigler-Hill & Monica, 2015). One such example is the HEXACO model of personality (Ashton & Lee, 2009), which builds on the existing big five model (comprising the aforementioned traits) but also proposes an additional sixth trait dimension of

honest/humility (see Table 14), which to date has not received much empirical attention in the context of video games. An explanation for why the sixth dimension is not a feature of the conventional big five perspective is related to the lexical hypothesis that underpins the big five perspective; that is, that major dimensions of personality are encoded in human language (Goldberg, 1993). However, more recent lexical studies have found support for a 6th factor in the Honesty-Humility dimension (Ashton & Lee, 2001). Importantly, the acknowledgement of the Honesty-Humility factor is not a new addition to the big five model per se, but rather is a repartitioning of existing factors (namely, variance between neuroticism and agreeableness). As such, Anglim and O'Connor (2019) recommend that from a predictive validity perspective, it is more appropriate to proceed with the inclusion of an addition factor (Honesty-Humility) than to not.

Table 14

Ashton and Lee (2004) HEXACO model of personality

Honesty-Humility	**Emotionality**	**Extraversion**	**Agreeableness**	**Conscientiousness**	**Openness to Experience**
H	**E**	**X**	**A**	**C**	**O**
Sincerity	Fearfulness, Anxiety	Social self-esteem	Forgivingness	Organisation	Aesthetic appreciation
Fairness	Anxiety	Social boldness	Gentleness	Diligence	Inquisitiveness
Greed Avoidance	Dependence	Sociability	Flexibility	Perfectionism	Creativity
Modesty	Sentimentality	Liveliness	Patience	Prudence	Unconventionality

The use of the HEXACO model is, however, contested in personality psychology, with some evidence that it does not replicate the underpinning model of the big five factor structure, especially when applied to other cultures (De Raad et al., 2010). Moreover, there is contention on whether any sixth component to the big five factor structure will likely correlate highly with one of the five factors, thereby becoming redundant (Saucier & Ostendorf, 1999). In the context of video games, the HEXACO model may prove a more effective measure of gamer personality due to the addition of the Honesty-Humility dimension. Studies have indicated that prosocial behaviour, such as cooperation and honesty, can emerge in video games (Greitemeyer & Osswald, 2009), indicating that the Honesty-Humility dimension can provide an additional and relevant point of personality

analysis, particularly in the context of gaming. Furthermore, prosocial gaming behaviour has been found to correlate with variance in user engagement (Gentile et al., 2009), which in the context of Gamification and its primary focus on improving user engagement, makes the HEXACO model and the sixth component it offers of arguably greater relevance when measuring gamer personality.

In relation to Gamification, the Honesty-Humility trait may also be highly useful. Given that Gamification is the application of design features to real-life contexts, it is reasonable to assume that users who are interacting within the gamified space will often be working closely with familiar users, such as colleagues in a workplace, students in a classroom, or members of a gym. As the Honesty-Humility dimension comprises of sub-components of fairness, sincerity, and greed avoidance, measuring how a user scores in these sub-traits could inform whether they are better suited or more receptive to forms of Gamification that require teamwork, or emphasise individual efforts.

As a domain of individual differences, personality more generally is likely to also influence user design preference, with research indicating that variance in trait dimensions is associated with variance in genre preference (or more specifically, design feature preference). To illustrate, Allam (2017) found that action games were more preferred by players with high trait neuroticism and low trait agreeableness, while genres which offer social and musical party features were more preferred by players with high trait extraversion (Peever et al., 2012). As such, personality provides another factor that can inform how Gamification can be better adapted to meet the needs of the end user and potentially increase engagement.

5.3 Rationale

The overarching aim of this PhD is to understand how user engagement with gamified systems/context can be tailored to meet the needs of the end user, by focusing on how individual differences, such as motivation and personality relate to user design preference. It is argued that by identifying how these individual differences can predict user design preference, Gamification can be effectively adapted to meet the needs of the user, thereby potentially maximising user engagement with the gamified context/system (in line with the primary purpose of Gamification).

Given that any identified relationships between individual differences and design feature preference will primarily inform how user engagement can be maximised via the design of Gamification, a preceding step in this study was to identify which design features, for which preference might vary, would predict changes in user engagement. For example, if an identified relationship between an individual difference and preference for a design feature would not predict variance in user engagement, in the context of informing Gamification design, such a relationship would not be useful, given that there is no indication that curating an experience based on the relationship would not necessarily result in a genuine increase of user engagement. As such, identifying which design features would predict variances in engagement was a key preliminary step.

Taking into consideration how the objectives of this study are situated within the broader, wider context of the overall project aims, Figure 5 illustrates the flow diagram. In summary, there are three elements of interest, all of which are related. The first is the individual difference of a user (such as their personality or motivation). The second is user design feature preference, which is expected to vary depending on the user's individual difference. Third is user engagement, which is expected to vary depending on the design feature preference the user possesses.

Figure 5 – Flow diagram illustrating role of individual differences

Several areas of improvement identified in the previous study were also addressed in the present study. In summary, there were four key considerations. The first, was to diversify the participant sample. In the previous study, players from only one gaming community

provided ratings of representativeness. According to Yee (2006), players are not a monolithic group, evident in the diversity of user motivation and personality, and the array of unique user experiences provided by a variety of game genres. The present study sought to include a broader spectrum of players by recruiting participants from multiple gaming communities. The second, was to increase the number of design feature vignettes from the original 37, given that further reviews revealed more design features that were suitable for inclusion into gamified contexts (e.g., a Punish design feature which provides users the ability to punish an opponent once you have defeated them).

The third, was to subject all design feature vignettes (from the previous study and the new additions in the present study) to a factor structure, such that the individual design feature vignettes could be more manageably categorised. For example, in assessing how design feature preference might relate to user motivation or personality, in its current form (of 37 design features) a quantitative approach to analysing any relationships might require multiple individual statistical tests (increasing a Type I error). In stratifying and effectively "reducing" the number of items into broader categories, such statistical obstacles can be avoided, as well as achieving model parsimony.

Notably, the objective to achieve model parsimony and a factor structure did not conflict with the broader objective to improve on issues of specificity outlined in section 4.2.3. A key component of the specificity issue concerns replicability of findings, such that future research can build accurately upon the design features used and tested in previous work. By reducing the number of design features into broader categories, this aspect of specificity is not compromised, given that broader factor categories and the individual design features with which the category is comprised, can still be clearly reported. Furthermore, researchers can be made aware via a style of reporting and classification that each factor should be regarded as an umbrella, under which the smaller, implementable design features have been operationalised and used within a given research study.

5.4 Research aim

The primary aim for the present study was to apply the DFPS to understand how user design preference related to dimensions of user motivation and personality, as measured by the Trojan Player Typology motivation scale (Kahn et al., 2015), and the HEXACO 60-item personality scale (Ashton & Lee, 2009). The secondary aims for the present study were to increase the variance in sample characteristics by recruiting from a greater number of gaming communities, expand the number of design features that comprise the DFPS from the existing 37, and to achieve a broader categorisation of the design features which comprise the DFPS.

5.5 Methods

5.5.1 Design

The present study employed a within-subjects cross-sectional online survey methodology to identify how motivation and personality relate to user design preference. Participants were asked to complete a user engagement scale, a motivation scale and personality scale, and provide preference ratings for a series of design feature vignettes. As such there were four variables of interest: reported user engagement as measured by play duration and play frequency questions; responses to the Trojan Player Typology motivation scale (Kahn et al., 2015) which measured user motivation; responses to the HEXACO-60 personality scale (Ashton & Lee, 2009) which measured user personality; and responses to the DFPS, which measured user preference for 47 design features.

5.5.2 Participants

Players were recruited from various gaming Reddits; /r/GameCollecting, /r/Games, /r/Gaming, /r/GoW, /r/Guildwars, /r/MachineLearning, /r/PCMR, /r/Rainbow6, r/Red Dead, /r/Rocket League, /r/Steam, /r/True Gaming, /r/Witcher, and /r/WoW. A total of 900 players participated (Reddit breakdown is detailed in Table 15), with 90% male, 7% female, and 3% not identifying. Ages ranged from 18 to 54 (Mean age = 25.55 years; SD = 5.32). The average level of gaming experience for the sample was 17 years (SD = 6.28). 56% of participants were employed, 29% were students, 10% were actively seeking employment, and 5% were unemployed. Of the total sample, 51% were from North America, 36% from Europe, 6% from Asia, and the remaining from South America,

Oceania, or Africa. 92% of participants reported most use of a PC platform to play video games, followed by consoles (XBOX, Playstation, Nintendo) at 7%, and mobiles and tablets at 1%. Most participants reported World of Warcraft as their favourite game (n=256), followed by DOTA 2 (n=22), the Final Fantasy series (n=21), the HALO series (n=20), and Rocket League (n=20).

Table 15

Reddit sample distribution (largest to smallest)

Reddit	***(n)***
WoW	500
Steam	123
PCMR	113
True Gaming	46
Witcher	34
Rocket League	30
Gaming	17
Red Dead	10
Rainbow 6	10
Game Collecting	8
GoW (Gears of War)	5
Games	2
Guildwars	1
Machine Learning	1

5.5.2.1 Eligibility criteria

Participants were only eligible to participate if they met two main criteria. The first, was that participants must have regularly played video games for at least two hours per week, which in previous gaming research has been regarded as the minimum time spent playing to qualify as being a gamer (Kolo & Braun, 2004). It was expected that players possess the prerequisite experience to understand the functional representation of Design features, therefore ratings of representativeness would be more reliable than those given by non-players. The second criteria were for participants to have been aged 18 years or older. As the study exclusively employed online methodologies, there was no way for researchers to obtain parental consent for underage participants, therefore only those of the legal age to consent were able to participate.

5.5.2.2 Participation incentive

All participants were automatically enrolled into a free game giveaway, wherein the successful winner would be awarded a game of their choice, limited to $77 or £55 and purchasable only from a reputable online seller (such as STEAM, Origin, or the XBOX marketplace).

5.5.3 Materials

5.5.3.1 Design feature preference

Participants were required to complete the DFPS, which required participants to report to what extent they found a given design feature fun, motivating, useful, and preferable (Lopez & Tucker, 2019). A total of 47 design features were functionally represented via vignettes (an increase of 10 from the previous study), each of which had been generated from a literature review (Arnab et al., 2015; da Rocha Seixa et al., 2016; Hamari & Lehdonvirta, 2010; Lameras, 2017; Marczewski, 2015; Nacke, 2018; Orji et al., 2018; Rocha et al., 2008; Sailer et al., 2017; Tondello et al., 2017; Werbach & Hunter, 2012), and previous study validation (see Table 16).

The additional 10 design features that were added to the previous 37 design features followed the same wording format as that which was found to be most representative in the previous study i.e., defining the design feature by wording using the possessive form. Repeated measures were also applied, such as each vignette being limited to 17 words and achieving an average Flesch reading ease score of 53.31 and a Flesch-Kincaid grade level test score of 8.85, suggesting that the vignettes could be easily read by 13-14-year-olds. Together, all design features achieved an alpha of (a = .934) suggesting high internal consistency. Ratings were made via a 5-point Likert scale (1 – strongly unfavourable, 5 – strongly favourable) that was designed to emulate a conventional star rating system to induce participants to consider their ratings in a more meaningful capacity (see Appendix C). Example vignettes include examples include the Complementarity design feature (being required to complete an objective or task only with the help and abilities of another player), and the Demotion design feature (being demoted and having your rank reduced after failing in some way).

5.5.3.2 User engagement

User engagement was measured by play duration (how long a typical play session lasts) and play frequency (how often a participant plays in a week), as has been used in previous gaming research (Brunborg et al., 2014). The conceptualisation of behavioural engagement in this study diverges from more technical understandings of behavioural engagement, which usually comprise of concepts such as Immersion, Presence, Flow, Psychological absorption, and Dissociation (Brockmyer et al., 2009). While these concepts can provide an insight into how engaged a user may be when playing video games, in the context of Gamification, they may not be appropriate.

For example, according to Brown and Cairns (2004), immersion is defined as a "psychological state in which a person's awareness of the physical self is diminished or lost due to being surrounded in an engrossing total environment" (p. 94). Given the “real-life” aspect of Gamification, it is highly unlikely that a loss of personal awareness due to environmental engrossment would be achievable, given the emphasis of being presently aware within the gamified context, and interacting with design features in a real-life setting. As such, determining which design features are most engaging using this metric is arguably inappropriate.

5.5.3.3 Motivation

Motivation was measured by the Trojan Player Typology Gaming Motivation Scale (Kahn et al., 2015), which requires participants to rate the extent to which they agree or disagree with 15 motivation related statements via a five-point likert scale (1- strongly disagree, 5 – strongly agree). The scale measures six dimensions of gaming motivation; Socialisers (players who play to socialise), Completionists (players who play to complete every aspect of the game), Competitors (players who play to succeed), Escapists (players who play to escape from real life), Story-driven (players who primarily are interested in the game story), and Smarty-pants (players who seek intellectual stimulation and challenge). Example statements include “I like to chat with my friends while playing a video game” and “It is important to me to be the fastest and most skilled person playing the game”.

5.5.3.4 Personality

Personality was measured by the HEXACO-60 (Ashton & Lee, 2009) personality test, which requires participants to rate the extent to which they agree with 60 statements, via a 5-point Likert scale (1- strongly disagree, 5 - strongly agree). The scale measures six dimensions of personality (10 items per dimension): Honesty-Humility (comprising of sub-traits such as sincerity, fairness, greed avoidance, and modesty), Emotionality (comprising of sub-traits such as fearfulness, anxiety, dependence, and sentimentality), Extraversion (comprising of sub-traits such as social self-esteem, social boldness, sociability, and liveliness), Agreeableness (comprising of sub-traits such as forgivingness, gentleness, flexibility, and patience), Conscientiousness (comprising of sub-traits such as organisation, diligence, perfectionism, and prudence), and Openness (comprising of sub-traits such as aesthetic appreciation, inquisitiveness, creativity, and unconventionality). Example statements include "People sometimes tell me that I am too critical of others" and "My attitude toward people who have treated me badly is "forgive and forget".

5.5.3.5 Demographics

Participants were asked to submit standard demographic data, such as age, ethnicity, gender, education status, country of domicile, employment status and marital status. In addition, general gaming related demographics will also be collected, such as favourite game title, favourite game genre to play, most used gaming platform, and for how long they have played video games (in years).

Table 16
All 47 design feature vignettes

Design feature	Vignette
PVP	Being able to compete against other players
Leaderboard (Competition)	Being able to see how your score and rank compare with other players
Punish	Being able to punish an opponent once you have defeated them
Complementarity	Being required to complete an objective or task only with the help and abilities of another player
Shared Goal	Being able to complete an objective more effectively if you work with another player
Trade	Being able to trade inventory items/currency with other players in exchange for items/currency
Friend Invite	Having the option to invite friends to play with you
Text Chat	Being able to communicate with other players via text chat

Voice Chat	Being able to communicate with other players via voice chat
Emotes	Being able to express your emotion and feelings through your avatar behaviour (such as jumping or dancing)
Communal discovery	Being able to complete tasks and learn new things as a community
Trophy	The scenario in which a trophy is received after completing an achievement or milestone
Badges	The scenario in which a badge is received after completing an achievement or milestone
Medal	The scenario in which a medal is received after completing an achievement or milestone
Tokens	Being able to earn tokens after completing challenges, that can be used to buy game content
Item Power-up	Being able to receive beneficial items / power ups after completing challenges or tasks
Bonnus	The scenario in which an unexpected or additional reward is received
Lottery	The scenario in which you receive a reward by luck
Depletion	Having to lose points, items or currency after failing in some way
Restriction	Having to lose access to some aspects of the game after failing in some way
Demotion	Being demoted and having your rank reduced after failing in some way
Points	Being able to see your progression in a number format
Progress Bar	Being able to see how close you are to reaching a milestone in a bar format
Leaderboard	Being able to see how your score and rank compare with other players
Scarlett letter	The element where other players are made aware of when you are stuck/failing
Performance graphs	Being able to view your performance level in a graph
Walkthrough	Having the option to receive a step-by-step guide on how to complete tasks or play the game
Tips / Hints	Being able to receive tips and hints when playing
Notification / Prompts	Being able to receive key notifications and updates when playing
Cut Scenes	Being able to see how the story progresses through cutscenes
Storyline	Having a central theme and story that the game revolves around
Currency	Being able to spend your in-game money/currency on game content
Item Degradation	Knowing that you must use some items/game content sooner rather than later due to time expiration
Dashboard	Being able to access game information, such as your game history, resources, profile, friends list, achievements etc.
Behavioural Momentum	The gradual increase in difficulty as the game goes on
Levels	Sections or parts of the game that are only accessible once a previous/existing level is completed
Barriers / Access	The exclusion from accessing specific aspects or parts of the game

Game Goal	The smaller and more immediate goals that once completed, will assist in fulfilling the game objective
Game Objective	The overarching goal when playing a game mode
Boss Battles	Overcoming an enemy/task at the end of a level that has increased significantly in difficulty
Beginners luck	Help in achieving a high rate of success when completing the first few tasks
Design / Editing / Customisation	The option to edit or design aspects of the game (e.g., avatar, environment, inventory)
Decision Making	The power to make decisions that significantly alter the course of the game/story
Vote	Having the opportunity to vote on something (e.g., map, weapon, rules)
Avatar	Being able to represent yourself via a virtual model/sprite/signature
Profile	Being able to immediately convey several aspects of yourself to other players
Rank / Status	Being assigned a category and rank that reflects your ability, score and/or experience

5.5.4 Procedure

Participants access the survey via a URL link provided in the recruitment advertisement, after which they were presented with the study information sheet. After providing consent and completing the demographics sheet, participants were required to report their level of engagement, followed by the completion of the Trojan motivation scale and the HEXACO personality scale. Thereafter, participants completed the design feature preference rating task, and were finally debriefed.

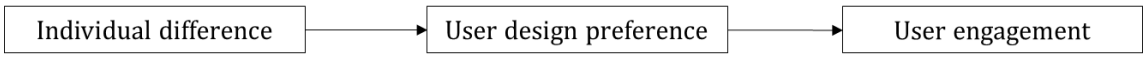
5.5.5 Ethics approval

Ethical approval was granted by Birmingham City University's research ethics committee under the reference code Lally/2024/R(A)/2019/Mar/BLSS.

5.6 Results

The analytical approach to data analysis for part one was spread across four analyses, in line with the research aims of the overall study. The first analysis was to stratify the 47 design feature vignettes, such that a more simplified and parsimonious model of design features could be reached. The second analysis was to identify which design features were predictive of variance in user engagement (play duration and play frequency) as based on user design feature preference. Analysis three and four were to focus on the role of individual differences in predicting user preference for significant design features, as per the path diagram below (Figure 5).

Figure 5 – Flow diagram illustrating role of individual differences



Specifically, analysis three focused on how user motivation predicted design feature preference, while analysis four focused on how user personality predicted design feature preference. Descriptive statistics showing mean scores for user engagement, user motivation, user personality, and design feature preference, as well as indications of normality (Skewness and Kurtosis), and the number of participants that provided responses is outlined in Table 34.

Table 34

Distribution statistics and mean scores for user engagement, user motivation, user personality, and design feature preference

Variable	Mean	SD	Skew (SE)	Kurt (SE)	*n*
User engagement					
Play duration	25.15	15.76	1.34 (.082)	2.45 (.163)	900
Play frequency	7.74	4.70	3.49 (.082)	20.11 (.163)	900
Motivation (Trojan Player Typology)					
Social motivation	11.28	2.97	-0.91 (.082)	0.26 (.163)	900
Completionist motivation	11.74	2.37	-0.71 (.082)	0.50 (.163)	900
Competitor motivation	9.37	3.04	-0.09 (.082)	-0.70 (.163)	900
Story motivation	5.28	1.93	0.23 (.082)	-0.33 (.163)	900
Smarty motivation	6.55	1.87	-0.22 (.082)	-0.18 (.163)	900
Escape motivation	7.32	1.88	-0.61 (.082)	0.00 (.163)	900
Personality (HEXACO)					
Honesty-Humility	34.05	6.17	-0.30 (.082)	0.36 (.163)	900
Emotionality	29.61	6.53	0.13 (.082)	-0.17 (.163)	900
Extraversion	28.31	7.55	-0.03 (.082)	-0.36 (.163)	900

Agreeableness	31.22	5.72	-0.20 (.082)	0.40 (.163)	900
Conscientiousness	34.13	6.08	-0.27 (.082)	-0.09 (.163)	900
Openness	35.17	6.44	-0.40 (.082)	-0.03 (.163)	900
Design features preference					
PVP	12.38	4.84	-0.34 (.082)	-0.66 (.163)	900
Leaderboard (Competition)	13.45	4.92	-0.54 (.082)	-0.53 (.163)	900
Punish	7.44	5.00	0.67 (.082)	-0.22 (.163)	900
Complementarity	10.44	5.32	0.04 (.082)	-0.96 (.163)	900
Shared Goal	13.98	4.67	-0.63 (.082)	-0.33 (.163)	900
Trade	14.43	4.49	-0.75 (.082)	0.15 (.163)	900
Friend Invite	17.13	3.96	-1.93 (.082)	3.89 (.163)	900
Text Chat	14.67	4.14	-0.87 (.082)	0.74 (.163)	900
Voice Chat	13.74	4.82	-0.62 (.082)	-0.35 (.163)	900
Emotes	10.53	5.18	-0.03 (.082)	-0.84 (.163)	900
Communal discovery	14.18	4.78	-0.79 (.082)	0.00 (.163)	900
Trophy	13.82	4.88	-0.67 (.082)	-0.27 (.163)	900
Badges	12.65	5.20	-0.39 (.082)	-0.75 (.163)	900
Medal	12.29	5.32	-0.33 (.082)	-0.79 (.163)	900
Tokens	14.08	5.22	-0.79 (.082)	-0.19 (.163)	900
Item Power-up	14.77	4.84	-0.95 (.082)	0.36 (.163)	900
Bonus	14.23	4.54	-0.68 (.082)	-0.01 (.163)	900
Lottery	6.07	4.32	0.89 (.082)	0.75 (.163)	900
Depletion	8.02	4.76	0.35 (.082)	-0.68 (.163)	900
Restriction	5.50	4.25	1.03 (.082)	0.96 (.163)	900
Demotion	8.28	4.90	0.30 (.082)	-0.71 (.163)	900
Points	14.86	4.16	-0.96 (.082)	0.75 (.163)	900
Progress Bar	15.34	4.04	-1.05 (.082)	0.91 (.163)	900
Leaderboard (Feedback)	13.71	5.09	-0.66 (.082)	-0.29 (.163)	900
Scarlett letter	8.01	5.03	0.37 (.082)	-0.63 (.163)	900
Performance graphs	13.73	5.01	-0.73 (.082)	-0.12 (.163)	900
Walkthrough	10.08	4.42	0.05 (.082)	-0.34 (.163)	900
Tips / Hints	11.23	4.30	-0.23 (.082)	-0.26 (.163)	900
Notification / Prompts	9.36	4.60	0.12 (.082)	-0.55 (.163)	900
Cut Scenes	14.76	4.62	-0.75 (.082)	-0.06 (.163)	900
Storyline	16.35	4.13	-1.36 (.082)	1.70 (.163)	900
Currency	15.27	4.54	-0.98 (.082)	0.38 (.163)	900
Item Degradation	6.87	4.50	0.59 (.082)	-0.11 (.163)	900
Dashboard	14.78	4.37	-0.78 (.082)	0.10 (.163)	900
Behavioural Momentum	16.23	4.10	-1.34 (.082)	1.68 (.163)	900
Levels	13.80	4.55	-0.65 (.082)	0.02 (.163)	900
Barriers / Access	7.96	4.94	0.35 (.082)	-0.61 (.163)	900
Game Goal	14.51	4.01	-0.71 (.082)	0.39 (.163)	900
Game Objective	15.83	3.91	-1.17 (.082)	1.71 (.163)	900
Boss Battles	16.77	3.77	-1.59 (.082)	2.91 (.163)	900
Beginners luck	8.86	4.88	0.20 (.082)	-0.61 (.163)	900
Design / Editing / Customisation	15.31	4.68	-1.04 (.082)	0.59 (.163)	900

Decision Making	16.16	4.25	-1.37 (.082)	1.71 (.163)	900
Vote	12.61	4.84	-0.47 (.082)	-0.27 (.163)	900
Avatar	14.08	5.07	-0.68 (.082)	-0.27 (.163)	900
Profile	12.26	5.23	-0.35 (.082)	-0.59 (.163)	900
Rank / Status	14.08	4.79	-0.73 (.082)	0.00 (.163)	900

Notes: SD = Standard deviation, Skew = Skewness, Kurt = Kurtosis, SE = Standard error, n = Sample size.

5.6.1 Analysis One – Exploratory Factor Analysis of design features

An EFA using a Principle-Axis Factor (PFA) extraction was conducted to identify the internal factor structure of 900 participant preference ratings to the 47 design feature vignettes that comprised the DFPS. The suitability of PAF was first assessed prior to analysis. Inspection of the correlation matrix showed that all variables had at least one correlational coefficient greater than 0.3 (Appendix D), highlighting that all variables (design features) are adequately correlated with each other. The overall Kaiser-Meyer-Olkin (KMO) measure was .919, with individual KMO measures all greater than 0.8 (Appendix E), further highlighting the linearity of relationships between each variable (design feature) and suitability for conducting the PAF (measures over .08 are considered highly suitable for factor analyses; Kaiser, 1974). Bartlett's Test of Sphericity was also statistically significant [$\chi 2$ (1081) = 19814.30, $p<0.01$], indicating that the correlation matrix did not suffer from multicollinearity – too high or too low correlations among all variables.

The PAF revealed nine components that had eigenvalues greater than 1 and which explained 24.9%, 5.7%, 4.3%, 3.7%, 3.1%, 2.6%, 2.4% 1.8% and 1.3% of the total variance, respectively. Visual inspection of the scree plot and identification of the inflection point indicated that nine components should be retained (Catell, 1966). In addition, a nine-component solution met the interpretability criterion, therefore nine components were retained. The nine-component solution explained 49.8% of the total variance. A Direct Oblimin Oblique rotation was used as factors (and constituent variables) were expected to be correlated (as per suitability tests discussed above). The rotated solution exhibited 'simple structure' (Thurstone, 1947). A coefficient suppression of 0.3 was desired (Tabachnick & Fidell, 2014), and factors which remained after this coefficient suppression were to have at least comprised of three loadings, each of which being greater than 0.4 (Samuels, 2016).

The interpretation of data was consistent with expectations of how each design feature would load onto factors, based on the similarity of each design feature function. For example, design features which served a function of creating, managing, or moderating the level of difficulty a user encounters loaded onto the first factor, referred to as the Difficulty factor (comprising of the design features; Behavioural Momentum (the gradual increase in difficulty as the game goes on), Levels (sections or parts of the game that are only accessible once a previous/existing level is completed), Game Goal (the smaller and more immediate goals that once completed will assist in fulfilling the game objective) Game Objective (the overarching goal when playing a game mode), and Boss Battle (overcoming an enemy/task at the end of a level that has increased significantly in difficulty). Similarly, Competition design features loaded onto Factor 2, Reward design features loaded onto Factor 3, Accessibility design features loaded onto Factor 4, Loss design features loaded onto Factor 5, Expression design features loaded onto Factor 6, Cooperation design features loaded onto Factor 7, Improvement design features loaded onto Factor 8, and Narrative design features loaded onto Factor 9. A detailed summary of which design feature each factor was comprised of can be found in Figure 6.

Figure 6 - Factor loadings from EFA

	Design feature dimension			
Design feature	**Difficulty (1)**	**Competition (2)**	**Reward (3)**	**Accessibility (4)**
	Promote increases in the level of difficulty experienced by the user	Design features which promote competition between users	Design features which promote the role of rewards available to users	Design features which promote user accessibility to the game
Behavioural momentum	0.510			
Levels	0.477			
Game Goal	0.453			
Game Objective	0.428			
Boss Battle	0.420			
Leaderboard (feedback)		0.683		
Leaderboard (Competition)		0.641		
Rank		0.584		
PVP		0.472		
Points		0.370		
Performance Graph		0.355		
Dashboard		0.342		
Badges			-1.002	
Medal			-0.966	

Trophy				-0.839	
Tips / Hints					0.759
Walkthrough					0.715
Beginners Luck					0.527
Notifications					0.521

	Design feature dimension (continued)				
	Loss (5)	**Expression (6)**	**Cooperation (7)**	**Improvement (8)**	**Narrative (9)**
Design feature	Design features which promote reminders of loss or failure to users	Design features which promote users to engage in personal expression	Design features which promote cooperative play between users	Design features which promote users to improve in-game skills or abilities	Design features which promote users to engage in story design and narrative
Restriction	-0.768				
Depletion	-0.693				
Demotion	-0.561				
Scarlett	-0.408				
Item Degradation	-0.394				
Barriers	-0.393				
Lottery	-0.336				
Punish	-0.313				
Avatar		0.825			
Design / Editing / Customisation		0.659			
Profile		0.609			
Vote		0.307			
Emotes		0.302			
Shared Goal			-0.764		
Complementarity			-0.649		
Communal Discovery			-0.520		
Voice Chat			-0.420		
Friend Invite			-0.408		
Text Chat			-0.394		
Item Power-up				0.603	
Currency				0.564	
Tokens				0.538	
Trade				0.506	
Progress Bar				0.344	
Storyline					0.882
Cut Scenes					0.606
Decision Making					0.506
Bonus					0.465

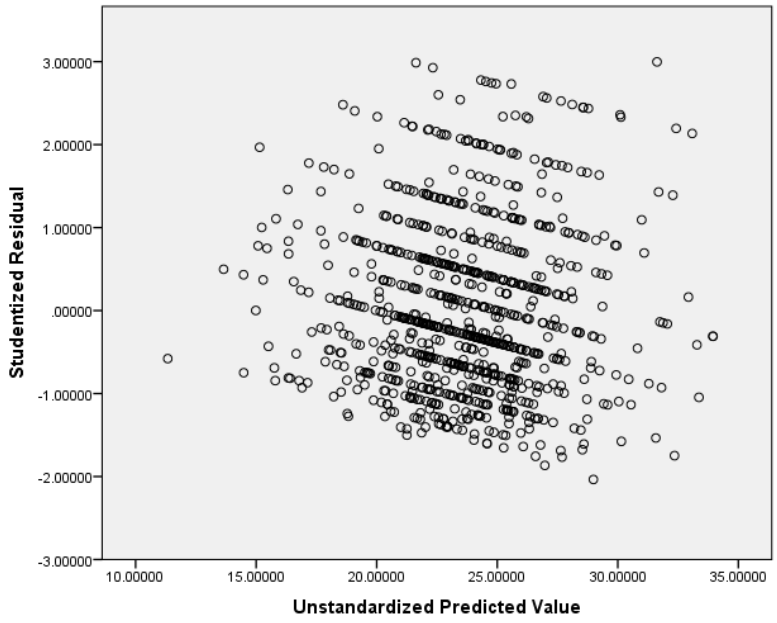
5.6.2 Analysis Two – Design feature preference and user engagement

5.6.2.1 Play duration

A multiple regression was conducted to assess how the identified nine dimensions of design feature preference predicted variance in play duration. There was linearity as assessed by partial regression plots and a plot of studentized residuals against predicted values. There was independence of residuals as assessed by a Durbin-Watson statistic of 1.785. There was homoscedasticity as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values (Figure 7). There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1 (Hair et al., 2014), indicating that the independent variables (design feature dimensions) were not highly correlated with each other.

Measures to detect outliers revealed no studentized deleted residuals greater than ±3 standard deviations (Everitt & Skrondal, 2010), no leverage values greater than 0.2 (Huber, 1981), or values (influential points) for Cook's distance above 1 (Cook, 1977), indicating no outliers were present in the data. The assumption of normality was met, as assessed by visual inspection of a Q-Q Plot (Appendix F). The multiple regression model significantly predicted play duration, $F(9, 863) = 5.966$, $p < .001$, adj. $R^2 = .05$. Only three of the nine design feature dimensions significantly predicted play duration: Accessibility Design features ($B = -.17$, $p < .001$), Expression Design features ($B = .09$, $p < .050$), and Improvement Design features ($B = .10$, $p < .050$). Whereas increases in preference for Expression and Improvement Design features correlated with an increase in play duration, preference for Accessibility Design features correlated with a decrease in play duration. The nonsignificant associations between the remaining six design feature dimensions and play duration can be found in Appendix X).

Figure 7 – Plot of play duration studentized residuals



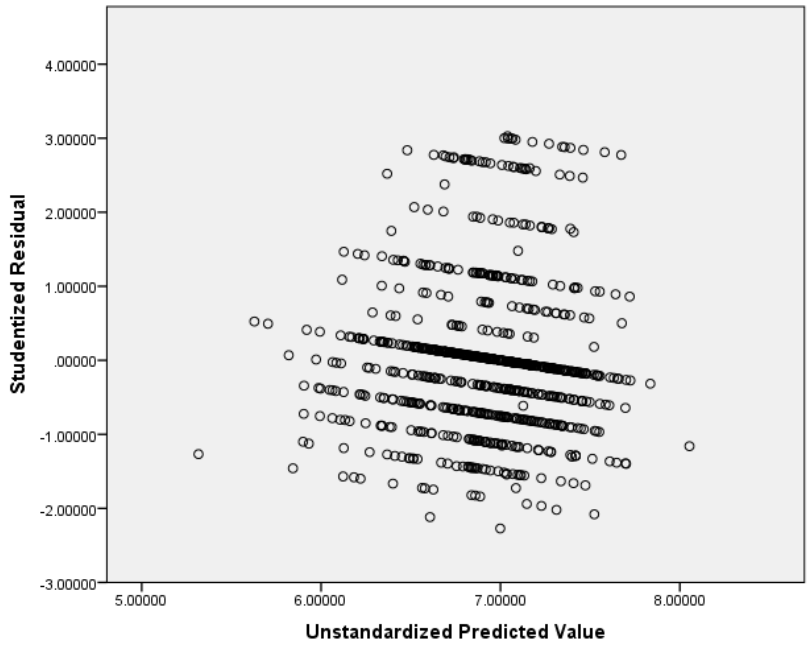
5.6.2.2 Play frequency

A multiple regression was conducted to assess how nine dimensions of design feature preference predicted variance in play frequency. There was linearity as assessed by partial regression plots and a plot of studentized residuals against predicted values. There was independence of residuals as assessed by a Durbin-Watson statistic of 1.991. There was homoscedasticity as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values (Figure 8). There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1 (Hair et al., 2014), indicating that the independent variables (design feature dimensions) were not highly correlated with each other.

Measures to detect outliers revealed no studentized deleted residuals greater than ±3 standard deviations (Everitt & Skrondal, 2010), no leverage values greater than 0.2 (Huber, 1981), or values (influential points) for Cook's distance above 1 (Cook, 1977), indicating no outliers were present in the data. The assumption of normality was met, as assessed by visual inspection of a Q-Q Plot (Appendix G). The multiple regression model significantly predicted play frequency, $F(9, 843) = 1.869$, $P < .050$, adj. $R^2 = .02$. Only one of nine design feature dimensions significantly predicted play frequency: Expression design features ($B = .02$, $p < .050$). Specifically, higher reported preference for Expression

design features corelated with increases in play frequency. The nonsignificant associations between the remaining eight design feature dimensions and play frequency can be found in Appendix X).

Figure 8 – Plot of play frequency studentized residuals



5.6.3 Analysis Three – User motivation and design feature preference

After determining which design features significantly predicted variance in user engagement (Expression, Improvement, and Accessibility design features) (EIA), analysis three focused on identifying how user motivation might predict variance in reported preference for EIA design features.

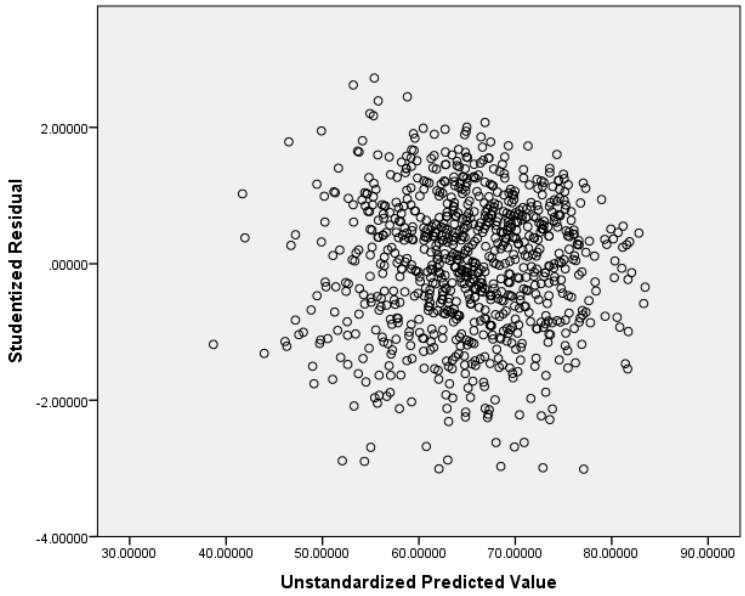
To ensure a Type I error did not occur due to the same test being conducted multiple times on the same data set, a Bonferroni correction was applied, such that the significance level for each of the succeeding analyses was made smaller (VanderWeele & Mathur, 2019). As per guidance from Napierala (2012), the significance level of 0.05 was reduced to a significance level of 0.006, using the Bonferroni formula of dividing the original α-value by the number of analyses on the dependent variable (0.05/8 = 0.006).

5.6.3.1 Expression design features

A multiple regression was conducted to assess how motivation dimensions from the Trojan Player Typology predicted variance in Expression design feature preference. There was linearity as assessed by partial regression plots and a plot of studentized residuals against predicted values. There was independence of residuals as assessed by a Durbin-Watson statistic of 1.968. There was homoscedasticity as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values (Figure 6). There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1 (Hair et al., 2014), indicating that the independent variables (motivation dimensions) were not highly correlated with each other.

Measures to detect outliers revealed no studentized deleted residuals greater than ±3 standard deviations (Everitt & Skrondal, 2010), no leverage values greater than 0.2 (Huber, 1981), or values (influential points) for Cook's distance above 1 (Cook, 1977), indicating no outliers were present in the data. The assumption of normality was met, as assessed by visual inspection of a Q-Q Plot (Appendix H). Using the Bonferroni adjusted significance level, the multiple regression model significantly predicted user design preference, $F(6, 879) = 30.646$, $P < .001$, adj. $R^2 = .17$. Preference for Expression design features was significantly predicted by the Socialiser motivation ($B = 1.14$, $p < .001$), Story-driven motivation ($B = 1.66$, $p < .001$), and Escapist motivation ($B = 2.31$, $p < .001$). Specifically, scoring higher in Socialiser, Story-driven, and Escapist motivations correlated with increases in preference for Expression Design features. The nonsignificant associations between the remaining three motivations and Expression design features can be found in Appendix Y).

Figure 9 – Plot of Expression design feature preference and motivation studentized residuals



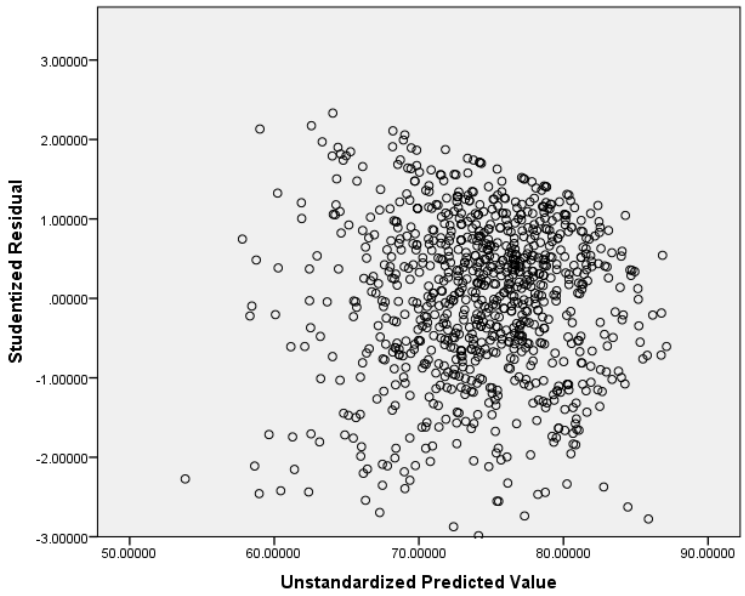
5.6.3.2 Improvement design features

A multiple regression was conducted to assess how motivation dimensions from the Trojan Player Typology predicted variance in Improvement design feature preference. There was linearity as assessed by partial regression plots and a plot of studentized residuals against predicted values. There was independence of residuals as assessed by a Durbin-Watson statistic of 1.968. There was homoscedasticity as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values (Figure 10). There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1 (Hair et al., 2014), indicating that the independent variables (motivation dimensions) were not highly correlated with each other.

Measures to detect outliers revealed no studentized deleted residuals greater than ±3 standard deviations (Everitt & Skrondal, 2010), no leverage values greater than 0.2 (Huber, 1981), or values (influential points) for Cook's distance above 1 (Cook, 1977), indicating no outliers were present in the data. The assumption of normality was met, as assessed by visual inspection of a Q-Q Plot (Appendix I). Using the Bonferroni adjusted significance level, the multiple regression model significantly predicted user design preference, $F(6, 877) = 18.364$, $P < .001$, adj. $R^2 = .11$. Preference for Improvement design features was significantly predicted by the Socialiser motivation ($B = 1.24$, $p < .001$), Story-driven motivation ($B = 1.26$, $p < .001$), and Escapist motivation ($B = .80$, $p < .050$).

Specifically, scoring higher in Socialiser, Story-driven, and Escapist motivations correlated with increases in preference for Improvement design features. The nonsignificant associations between the remaining three motivations and Improvement design features can be found in Appendix Y).

Figure 10 – Plot of Improvement design feature preference and motivation studentized residuals



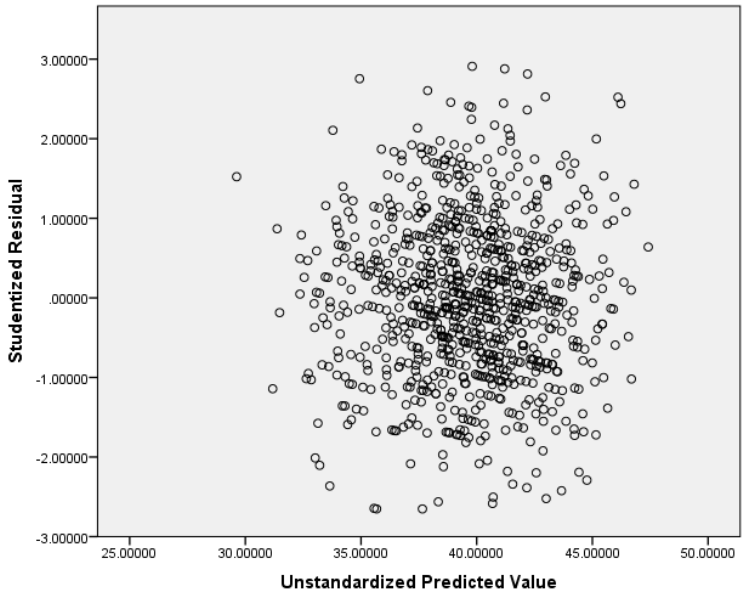
5.6.3.3 Accessibility design features

A multiple regression was conducted to assess how motivation dimensions from the Trojan Player Typology predicted variance in Accessibility design feature preference. There was linearity as assessed by partial regression plots and a plot of studentized residuals against predicted values. There was independence of residuals as assessed by a Durbin-Watson statistic of 2.072. There was homoscedasticity as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values (Figure 8). There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1 (Hair et al., 2014), indicating that the independent variables (motivation dimensions) were not highly correlated with each other.

Measures to detect outliers revealed no studentized deleted residuals greater than ±3 standard deviations (Everitt & Skrondal, 2010), no leverage values greater than 0.2 (Huber, 1981), or values (influential points) for Cook's distance above 1 (Cook, 1977), indicating no outliers were present in the data. The assumption of normality was met, as assessed by visual inspection of a Q-Q Plot (Appendix I). Using the Bonferroni

adjusted significance level, the multiple regression model significantly predicted user design preference, F (6, 881) = 7.193, p < .001, adj. R^2 = .04. Preference for Accessibility design features was significantly predicted by the Story-driven motivation (B = .77, p < .050), and Escapist motivation (B = .83, p < .050). Specifically, scoring higher in Story-driven and Escapist motivations correlated with increases in preference for Accessibility design features. The nonsignificant associations between the remaining four motivations and Accessibility design features can be found in Appendix Y).

Figure 11 – Plot of Accessibility design feature preference and motivation studentized residuals



5.6.4 Analysis Four – User personality and design feature preference

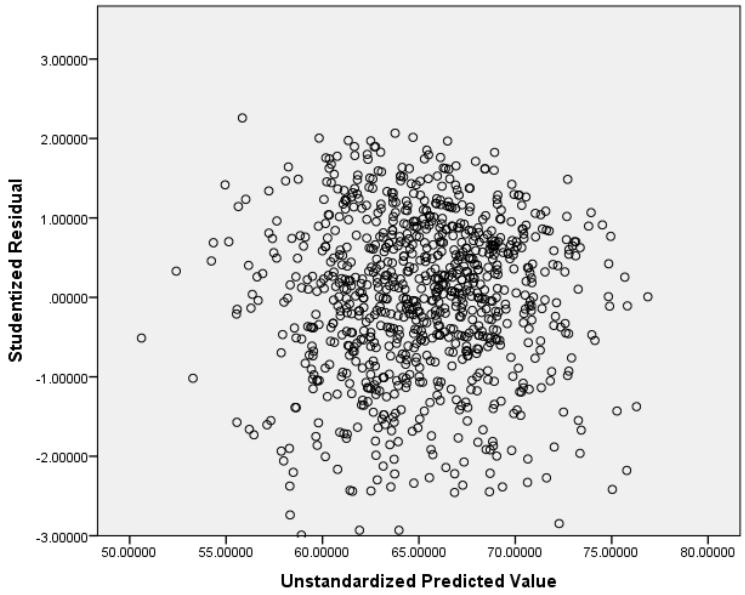
Analysis four follows analysis three, by focusing on how preference for EIA design features could be predicted by the individual difference of personality. As was the case for tests conducted in the previous section (5.6.3) on user motivation and design feature preference, all tests were subject to the Bonferroni adjusted significance level of 0.006, so as to reduce the occurrence of a Type I error.

5.6.4.1 Expression design features

A multiple regression was conducted to assess how personality dimensions from the HEXACO personality model predicted variance in Expression design feature preference. There was linearity as assessed by partial regression plots and a plot of studentized residuals against predicted values. There was independence of residuals as assessed by a Durbin-Watson statistic of 1.825. There was homoscedasticity as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values (Figure 12). There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1 (Hair et al., 2014), indicating that the independent variables (personality dimensions) were not highly correlated with each other.

Measures to detect outliers revealed no studentized deleted residuals greater than ±3 standard deviations (Everitt & Skrondal, 2010), no leverage values greater than 0.2 (Huber, 1981), or values (influential points) for Cook's distance above 1 (Cook, 1977), indicating no outliers were present in the data. The assumption of normality was met, as assessed by visual inspection of a Q-Q Plot (Appendix K). Using the Bonferroni adjusted significance level, the multiple regression model significantly predicted user design preference, $F(6, 877) = 8.486$, $p < .001$, adj. $R^2 = .05$. Preference for Expression design features was significantly predicted by personality dimensions of Emotionality ($B = .188$, $p < .001$), Agreeableness ($B = .084$, $p < .050$), and Openness ($B = .092$, $p < .050$). Specifically, scoring higher in personality dimensions of Emotionality, Agreeableness, and Openness correlated with increases in preference for Expression design features. The nonsignificant associations between the remaining three personality traits and Expression design features can be found in Appendix Z).

Figure 12 - Plot of Expression design feature preference and personality studentized residuals



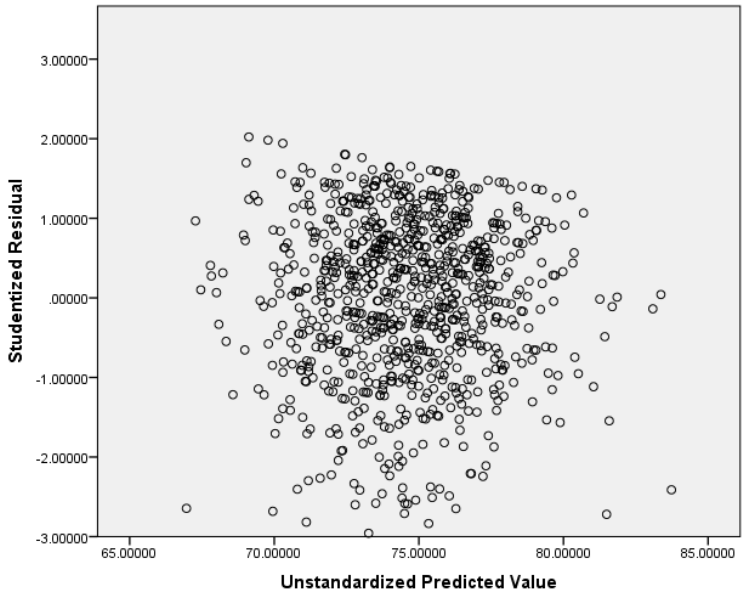
5.6.4.2 Improvement design features

A multiple regression was conducted to assess how personality dimensions from the HEXACO personality model predicted variance in Improvement design feature preference. There was linearity as assessed by partial regression plots and a plot of studentized residuals against predicted values. There was independence of residuals as assessed by a Durbin-Watson statistic of 1.963. There was homoscedasticity as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values (Figure 13). There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1 (Hair et al., 2014), indicating that the independent variables (personality dimensions) were not highly correlated with each other.

Measures to detect outliers revealed no studentized deleted residuals greater than ±3 standard deviations (Everitt & Skrondal, 2010), no leverage values greater than 0.2 (Huber, 1981), or values (influential points) for Cook's distance above 1 (Cook, 1977), indicating no outliers were present in the data. The assumption of normality was met, as assessed by visual inspection of a Q-Q Plot (Appendix L). Using the Bonferroni adjusted significance level, the multiple regression model significantly predicted user design preference, $F(6, 881) = 4.166$, $p < .001$, adj. $R^2 = .02$. Preference for Improvement design

features was significantly predicted by personality dimensions of Extraversion (B = .118, p < .050) and Emotionality (B = .097, p < .050). Specifically, scoring higher in personality dimensions of Extraversion and Emotionality correlated with increases in preference for Improvement design features. The nonsignificant associations between the remaining four personality traits and Improvement design features can be found in Appendix Z).

Figure 13 - Plot of Improvement design feature preference and personality studentized residuals



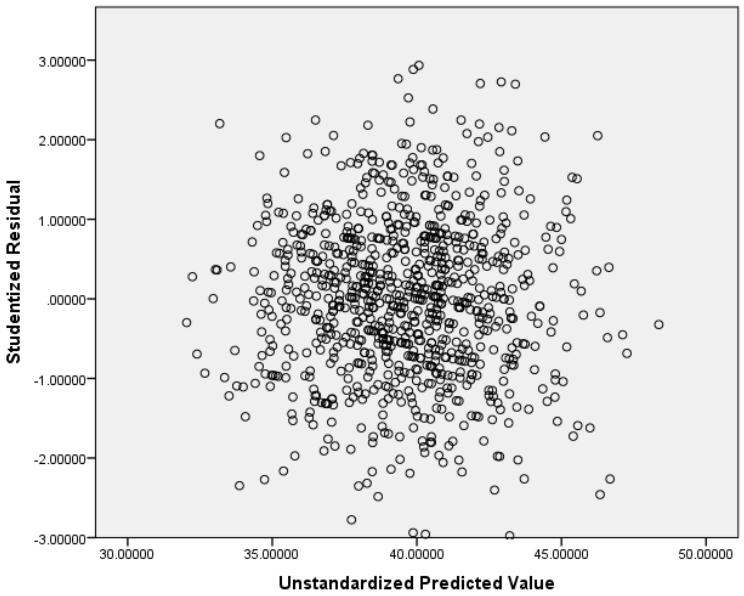
5.6.4.3 Accessibility design features

A multiple regression was conducted to assess how personality dimensions from the HEXACO personality model predicted variance in Accessibility design feature preference. There was linearity as assessed by partial regression plots and a plot of studentized residuals against predicted values. There was independence of residuals as assessed by a Durbin-Watson statistic of 2.051. There was homoscedasticity as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values (Figure 14). There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1 (Hair et al., 2014), indicating that the independent variables (personality dimensions) were not highly correlated with each other.

Measures to detect outliers revealed no studentized deleted residuals greater than ±3 standard deviations (Everitt & Skrondal, 2010), no leverage values greater than 0.2 (Huber, 1981), or values (influential points) for Cook's distance above 1 (Cook, 1977), indicating no outliers were present in the data. The assumption of normality was met, as

assessed by visual inspection of a Q-Q Plot (Appendix M). Using the Bonferroni adjusted significance level, the multiple regression model significantly predicted user design preference, $F(8, 892) = 5.837$, $p < .001$, adj. $R^2 = .03$. Preference for Accessibility design features was significantly predicted by personality dimensions of Agreeableness ($B = .079$, $p < .050$), Emotionality ($B = .160$, $p < .001$), Extraversion ($B = .087$, $p < .050$), and Honesty-Humility ($B = -.093$, $p < .050$). Specifically, scoring higher in the personality dimensions of Agreeableness, Emotionality, and Extraversion correlated with increases in preference for Accessibility design features, whereas higher scores in Honesty-Humility corresponded to decreases in preference for Accessibility design features. The nonsignificant associations between the remaining two personality traits and Accessibility design features can be found in Appendix Z).

Figure 14 - Plot of Accessibility design feature preference and personality studentized residuals



5.7 Discussion

The primary aim of the present study was to understand how design feature preference, as measured by the DFPS, related to user engagement, and how variances in user engagement and design feature preference could be predicted by the individual differences of motivation and personality. This aim was fulfilled, with the findings first indicating that variance in user engagement could be predicted by preference for Expression (positive association), Improvement (positive association), and Accessibility (negative association) design feature dimensions. As such, the subsequent analyses on how design feature preference related to user motivation and user personality, were focused on these dimensions. User engagement was not predicted by preference for the Difficulty, Competition, Reward, Loss, Cooperation, and Narrative design feature dimensions, each of which returned a non-significant association (Appendix X).

In examining how variance in design feature preference could be predicted by the individual difference of motivation, results indicated that the Trojan motivations of Socialiser, Story-driven, and Escapist predicted variance in preference for the EIA design features, whereas the motives of Completionist, Competitors, and Smarty-pants did not predict variances in preference for the EIA design features (Appendix Y). Specifically, preference for Expression design features was predicted by variance in the motivations of Socialiser, Story-driven, and Escapist; preference for Improvement design features was predicted by variance in the motivations of Socialiser, Story-driven, and Escapist; and preference for Accessibility design features was predicted by variance in the motivations of the Story-driven and Escapist.

In examining how variance in design feature preference could be predicted by the individual difference of personality, results indicated that the HEXACO personality traits of Agreeableness, Emotionality, Honest/Humility, Extraversion, and Openness to Experience, predicted variance in preference for EIA design features, whereas the personality trait of Conscientiousness did not predict preference for EIA design features (Appendix Z). Specifically, preference for Expression design features was predicted by variance in the personality dimensions of Emotionality, Agreeableness, and Openness to Experience; preference for Improvement design features was predicted by variance in the personality dimensions of Extraversion and Emotionality; and preference for

Accessibility design features was predicted by variance in the personality dimensions of Honesty-Humility, Emotionality, Extraversion and Agreeableness

Secondary aims of the present study were to expand the number of design features that comprise the DFPS, to subject the DFPS to factor structuring, and to diversify the participant sample. These aims were fulfilled, with the number of design features that comprise the DFPS increasing from 37 to 47. Additionally, the data collected from the DFPS was subject to factor structuring via an EFA, which revealed a nine-component factor structure, comprising of the design feature dimensions of Difficulty (1) Competition (2) Reward (3) Accessibility (4) Loss (5) Expression (6) Cooperation (7) Improvement (8), and Narrative (9). The DFPS structure is a novel contribution to the field of Adaptive Gamification and enables further model development. A more diverse participant sample was also used in the present study, with data collected from a total of 14 gaming communities (compared with the single gaming community used in the previous study).

5.7.1 User engagement and design feature preference

Before assessing the relationship between user design preference and individual differences, it was first necessary to identify how user design preference was related to engagement. As engagement is the primary aspect of the user experience that Gamification seeks to increase, it was necessary to focus on the design features for which variance in user preference was associated with variance in user engagement. In reference to the flow diagram illustrated earlier in this chapter, the focus on user design feature preference and engagement focuses on the latter part of the diagram (Figure 15).

Figure 15 - Flow diagram illustrating focus of analysis

Results indicated that EIA design features (detailed in Table 17) were found to predict variance in user engagement. Specifically, increased preference for Expression design features corresponded to both increased play duration and play frequency, and increased preference for Improvement design features corresponded to increased play duration. In contrast, increased preference for Accessibility design features corresponded to an overall decrease in play duration.

Table 17

Design feature dimensions that predicted variance in user engagement

Accessibility	Expression	Improvement
Tips / Hints	Avatar	Item Power-up
Being able to receive tips and hints when playing	Being able to represent yourself via a virtual model/sprite/signature	Being able to receive beneficial items / power ups after completing challenges or tasks
Walkthrough	Design / Editing / Customisation	Currency
Having the option to receive a step-by-step guide on how to complete tasks or play the game	The option to edit or design aspects of the game (e.g., avatar, environment, inventory)	Being able to spend your in-game money/currency on game content
Beginners Luck	Profile	Tokens
Help in achieving a high rate of success when completing the first few tasks	Being able to immediately convey several aspects of yourself to other players	Being able to earn tokens after completing challenges, that can be used to buy game content
Notifications	Vote	Trade
Being able to receive key notifications and updates when playing	Having the opportunity to vote on something (e.g., map, weapon, rules)	Being able to trade inventory items/currency with other players in exchange for items/currency
	Emotes	Progress Bar
	Being able to express your emotion and feelings through your avatar behaviour (such as jumping or dancing)	Being able to see how close you are to reaching a milestone in a bar format

The overall metric of engagement comprised of play frequency (how often a participant plays in a week) and play duration (how long a typical play session lasts). Similar measures have been employed in previous gaming research (Brunborg et al., 2014), and provide a more objective (quantified) metric than more sophisticated concepts of engagement (such as flow, immersion, presence, psychological absorption, and dissociation; Brockmyer et al., 2009) that rely on more subjective self-report data that focus on user experience (e.g., "I felt involved in this shopping task" or "This shopping experience was fun"; O'Brien, 2010). As such, the use of this metric was thought to be well suited to the aims of the study.

5.7.1.1 Expression dimension preference and engagement

The positive relationship between preference and user engagement (duration/frequency) could be explained by the nature of Expression design features. For example, the Avatar and Design / Editing / Customisation design features have what

could be described as significant replay value (i.e., they are features that maintain interactivity despite repeated user exposure or interaction). As these features facilitate and encourage user creativity and self-expression, it could be argued that a greater preference for these design features could indicate a greater desire on behalf of the user to create and express, therefore leading to increases in user engagement. Existing research also supports this position, with studies indicating that the ability to customise and edit can often lead to heightened user engagement (Leigh et al., 2018). Moreover, the role played by game publishers could also explain why increased preference corresponded to increased engagement. Modern games can often make available expansion packs, in-game purchases, and lootboxes, all of which can provide users with additional points of customisation, such as more character skins or game maps. On the basis that a greater range of customisation could lend itself to increases in user engagement (Macey & Hamari, 2018; Shibuya et al., 2019), it could be interpreted that greater customisability can lead to increased user engagement.

Additionally, Expression design features more broadly serve the function of creating an online (or in-game) identity. Take for instance MMORPGs, which offer a strong component of role-play. Excessive engagement (in the form of internet gaming addiction) is often observed in users of these games due to the facility to escape, roleplay and reinvent oneself (Kuss, 2013). Therefore, the notion of experimenting or reconstructing one's identity, across the way in which they appear to other users (Avatar), the way they interact with other users (Emotes) and the way in which they make decisions (Votes) draw attention to a form of identity fluidity that Expression design features facilitate, which could also explain why increased preference predicted increased engagement.

5.7.1.2 Improvement dimension preference and engagement

Mutual increases between preference for Improvement design features and user engagement (specifically play duration) also emerged. Play duration may increase along with preference for the design feature of Items / Powerups, as users could extend their time spent playing to make use of a newly acquired ability. For example, in a racing game, after receiving car enhancements, a user may spend longer playing the game to "try out" their new abilities. According to Denisova and Cook (2019), users who acquire enhancements tend to feel more immersed in a game; and as high immersion is a strong

predictor of increased user engagement (Seah & Cairns, 2008), it could be expected that the acquisition of such enhancements can lead to users increasing play duration.

Moreover, if a game requires users to spend in-game points or tokens to purchase enhancements, then a preference for the Currency design feature might also explain increased play duration, given that users will need to complete more in-game tasks as a way to accrue more currency (that can be spent on enhancements). Supporting studies also highlight the role of in-game transactions as spurring increased engagement (King et al., 2019; Nenad, 2017). Other Improvement design features, such as the Progress Bar design feature could also provide insight into why a mutual increase between preference and engagement was observed. The Progress Bar design feature serves the function of users being able to visually see how close they are to fulfilling a milestone. If interacting with this design feature, users may increase overall play duration to fulfil the Progress Bar and achieve the represented milestone, thereby an increased preference of the Progress Bar design feature could lead to increased user engagement.

5.7.1.3 Accessibility dimension preference and engagement

Unlike Expression and Improvement design features, the results indicated that increased preference for Accessibility design features corresponded to a decrease in user engagement. The direction of this relationship might be explained by how Accessibility design features are presented in-game. Take for example the Walkthrough and Beginners Luck design features, both of which aim to equip new users with the knowledge on how to perform actions and complete in-game tasks as part of a game's onboarding process. As a user progresses through initial tutorials and demonstrates a level of competence, the game will gradually reduce the presence and availability of these design features as the underlying function they serve (i.e., to "train" players) has been fulfilled and is no longer necessary (White, 2014). If a user prefers these design features, but the game gradually restricts or reduces the extent to which a user can interact with these design features, then a decrease in play duration would be expected.

Beyond the function of Accessibility design features, the inverse relationship found in the present study might also be explained by game-specific factors. For example, Andersen et al. (2012) note that the perceived usefulness of game tutorials from a player perspective depends largely on game complexity. That is, tutorials for games that are

more sophisticated and cognitively demanding are more favourably perceived than tutorials for games that are simpler. One explanation for this might be that users require a greater level of support in learning how to operate the game if the game is more complex. In line with the previous point about the rate at which a user becomes familiar with a game, tutorials for more complex games may be interacted with for longer periods given the function they serve is not as readily fulfilled as would be the case for less complex games. Support for this perspective comes with research indicating that tutorials of a more complex nature increase play time to a much greater extent than tutorials of less complex games (Andersen et al., 2012).

As such, if user engagement with Accessibility design features is based upon their perceived value or usefulness, and if perceived value or usefulness is largely influenced by the complexity of the video game (with tutorials of more complex games being more useful), then the observed relationship could be explained by the participants own gaming experience. That is, their experience with interacting with Accessibility design features might be characterised by either tutorials for games that were not complex, or for games where the user was already familiar and did not require the tutorial (thus a lower perceived usefulness of the tutorials), hence a lower reported play duration.

5.7.2 Individual differences and design feature preference

The next stage of analysis was to identify how individual differences of user motivation and personality could predict preference for the EIA design features. The focus on motivation, personality, and user design preference focuses on the former part of Figure 16.

Figure 16 - Flow diagram illustrating focus of analysis

5.7.2.1 Motivation and Expression dimension preference

Preference for Expression design features (Table 17) was predicted by the Socialiser, Story-driven, and Escapist (Table 13) motives, with increased motivation corresponding to increased preference. According to Kahn et al. (2015), users who foster a Socialiser motivation would be expected to have more in-game social relationships, which

complement an increased preference for design features that require a social context, such as the Emotes design feature. Similarly, Escapist motives, which are characterised by user's being drawn to fantasy and in-game power, complement an increased preference for characterised by users' being drawn to fantasy (Sherry et al., 2006) complement an increased preference for the Avatar, Design / Editing / Customisation, and Profile design features, all off which facilitate the user to role-play and virtualise aspects of themselves (and in many cases changes those aspects).

Table 13

Kahn et al. (2015) Trojan Player Typology

Socialisers	**Completionists**	**Competitors**	**Escapists**	**Story-driven**	**Smarty-pants**
Play to socialise	Play to complete every aspect of the game	Play to succeed	Play to escape from real life	Interested in the game story/narrative	Play for intellectual stimulation and challenge

Table 17

Design feature dimensions that predicted variance in user engagement

Accessibility	**Expression**	**Improvement**
Tips / Hints	Avatar	Item Power-up
Being able to receive tips and hints when playing	Being able to represent yourself via a virtual model/sprite/signature	Being able to receive beneficial items / power ups after completing challenges or tasks
Walkthrough	Design / Editing / Customisation	Currency
Having the option to receive a step-by-step guide on how to complete tasks or play the game	The option to edit or design aspects of the game (e.g., avatar, environment, inventory)	Being able to spend your in-game money/currency on game content
Beginners Luck	Profile	Tokens
Help in achieving a high rate of success when completing the first few tasks	Being able to immediately convey several aspects of yourself to other players	Being able to earn tokens after completing challenges, that can be used to buy game content
Notifications	Vote	Trade
Being able to receive key notifications and updates when playing	Having the opportunity to vote on something (e.g., map, weapon, rules)	Being able to trade inventory items/currency with other players in exchange for items/currency
	Emotes	Progress Bar
	Being able to express your emotion and feelings through your avatar behaviour (such as jumping or dancing)	Being able to see how close you are to reaching a milestone in a bar format

5.7.2.2 Motivation and Improvement dimension preference

Preference for Improvement design features (Table 17) was predicted by the Socialiser, Story-driven, and Escapist motives (Table 13), with increased motivation corresponding to increased preference. According to Chang et al. (2018), Escapist motives can be characterised by a desire to disconnect from real life, with an orientation towards achieving states of immersion and engaging in fantasy. Research has demonstrated that the achievement of such states is often accompanied by prolonged periods of play, meaning that to become immersed and escape the trials of real life, play duration increases (Michailidis et al., 2018). As such, a motive to disengage with reality and engage with aspects of fantasy could consist with a preference for Improvement design features which enable or are related to aspects of fantasy. For example, the Item Power-up design feature serves the function of providing users with character enhancements. The motive of fantasy in video games has been described as the ability to do things you cannot do in real life (Sherry et al., 2006) and stepping out of one's usual identity to adopt or try a new identity (Dmetrovics et al., 2011), thus design features which can provide enhancement of fantasy characters or identities would likely be preferred by those who are motivated to Escape. Moreover, design features which enable continuous development of such fantasy characters and identities might also be expected to be preferred by users who express Escape motives. For instance, the Improvement design features of Trade and Currency can be used to acquire content that can be applied to fantasy characters and identities (such as character skins; Macey & Hamari, 2018).

5.7.2.3 Motivation and Accessibility dimension preference

Preference for Accessibility design features (Table 17) was predicted by the Story-driven and Escapist motives (Table 13), with increased motivation corresponding to increased preference. As the Story-driven motive is characterised by an interest in the game narrative, the relationship could be explained when considering that the Walkthrough design feature might fulfil the motives of users who are Story-driven. To illustrate, as part of introductory tutorials and onboarding processes, game designers often imbed task completion and learning within the story narrative or game context (White, 2014), such that the user becomes familiar with the game controls and instructions, as well as understanding the world within which their activity will take place. Thus, it could be interpreted that a Walkthrough design feature might possess a sufficient degree of

narrative and game story and provide an environment in which the Story-driven motive is fulfilled. Additionally, the Tips / Hints design feature might also provide a user with exposure to the game narrative and story. For example, game tips and hints can often be conveyed during a loading screen (a small delay that occurs while a new level or segment of a game loads in the background, much like how online videos may buffer before being played). Often, the content included in loading screens can contain important information ranging from game lore and narrative to technical tips on how to operate more effectively in the game.

Escapist motives also consist with a preference for Accessibility design features. As discussed, Escapist motives are closely related to the notion of user immersion and flow (Kahn et al., 2015) wherein user attention and awareness is fully absorbed into the game world. A key facilitator to achieving the psychological state of immersion or flow is when a user reached an optimal balance of demand and competence, such that the task or obstacle they face is met with a level of ability and skill. In the context of users who foster an Escapist motive, honing their in-game ability and skill is likely to be supported by the Accessibility design features. For instance, the Tips / Hints design features could provide users with an understanding of how to optimise their skill development or overcome an obstacle. Additionally, the Walkthrough design feature would likely assist users in becoming familiar and accustomed to the game world, further supporting the development of their skill and ability. As such, the preference of Accessibility design features being associated with a motive of Escapism can be understood from the perspective that Accessibility design features support the development of one's competence to overcome in-game obstacles.

5.7.2.4 Personality and Expression dimension preference

Preference for Expression design features (Table 17) was predicted by the personality dimensions of Emotionality, Agreeableness, and Openness to Experience (Table 14). These findings can be explained by how the functions the Expression design feature serve can complement the trait tendencies of the personality dimensions. For instance, those who score highly in Emotionality are likely to experience anxiety in response to life stress, and fear danger (Ashton & Lee, 2004). In the context of preferring Expression design features, a high susceptibility to anxiety and sensitivity to danger could prompt a user to seek a safer context in the virtual world. This tendency is supported by the Preference for

Online Social Interaction theory, which posits that users, often those who experience social anxiety (Marino et al., 2020), believe that one is safer, more confident, and more comfortable when interacting virtually than when interacting in traditional social activities (Caplan, 2003). On this basis, it would appear consistent for variance in Emotionality to predict preference for Expression design features, particularly those that enable users to express themselves via an online identity (Besseiere et al., 2007), and facilitate a reinvention of oneself online, such as the Avatar design feature. For example, research highlights that an Avatar serves as a virtual representation of oneself (Gaetan et al., 2012) that can embody several aspects of the user's real and artificial self (Craipeau & Seys, 2005).

Table 14

Ashton and Lee (2004) HEXACO model of personality

Honesty-Humility	**Emotionality**	**Extraversion**	**Agreeableness**	**Conscientiousness**	**Openness to Experience**
H	**E**	**X**	**A**	**C**	**O**
Sincerity	Fearfulness, Anxiety	Social self-esteem	Forgivingness	Organisation	Aesthetic appreciation
Fairness	Anxiety	Social boldness	Gentleness	Diligence	Inquisitiveness
Greed Avoidance	Dependence	Sociability	Flexibility	Perfectionism	Creativity
Modesty	Sentimentality	Liveliness	Patience	Prudence	Unconventionality

In the case of Openness, users who score high in this personality dimension possess strong interest in art, nature, and make frequent use of their imagination (Ashton & Lee, 2004). In relation to user design preference, one could consider Expression design features, particularly the Avatar, Design / Editing / Customisation, and Emotes design features to be intuitive extensions and suitable features with which high Openness can be complemented and nurtured (San-Martin et al., 2020). For example, the Avatar design feature serves the function of allowing users to create a visual representation of oneself, while the Design / Editing / Customisation design features serves the function of providing users with a degree of control over how the virtual world can be changed. In both cases, the free use of imagination is well facilitated, therefore a user who scores high in Openness would be expected to prefer such design features.

Those who score high in Agreeableness are likely to be more accepting and lenient in the judgement of others, less competitive, and instead seek more cooperative play (Ferguson et al., 2020; Ross et al., 2002). As Expression design features tend to be more present and emphasised in game genres which also aim to provide more cooperative and collective user experiences, such as MMORPGs (e.g., World of Warcraft, Guild Wars, Elder Scrolls: Skyrim), users who score high in Agreeableness may prefer design features that are characteristic of these genres. One such example is the Expression design feature of Design / Editing / Customisation, and Avatar. In addition, as users who score high in Agreeableness are less likely to judge others, one could also expect Expression design features as being more preferable, given that the perceived judgement by others can inhibit self-expression (Bailey et al., 2020).

Building on this argument, the role of gender may also provide some explanation as to why increased Agreeableness predicted increased preference for Expression design features. Several studies indicate that gender differences exist in personality and temperament, with high Agreeableness observed in females more than males (Braakmann, 2009; Chapman et al., 2007; Mueller and Plug, 2006; Nyhus & Pons, 2012; Weisberg et al., 2011;). In the present study, of the 7% of females who participated (n = 66), 50 % (n = 30) reported a MMORPG as being their favourite game, further highlighting that increased Agreeableness can be associated with increased preference for game genres or worlds wherein there is an increased emphasis and presence of Expression design features.

5.7.2.5 Personality and Improvement dimension preference

Preference for Improvement design features (Table 17) was predicted by the personality dimensions of Extraversion and Emotionality (Table 14). As described previously, users who score high in Emotionality are likely to find the artificial nature of the virtual world as a safer hedge to manage their high sensitivity to danger and anxiety. In the context of these individuals preferring Improvement design features, it could be argued that the pursuit of a goal-directed activity, such as improving one's in-game competence, would be a safer arena in which difficulty can be faced, as opposed to pursuing such forms of improvement in real life. To elaborate, take for example the Tokens design feature, which serves the function of enabling users to earn tokens after completing challenging that can be used to buy game content. Though challenges can vary in the way they materialise, it

would be reasonable to assume that the level of difficulty intrinsic to each challenge will increase along with the reward that the user will receive if the challenge is completed. A user who scores high in Emotionality may perceive the process of facing difficulty to acquire the means to make purchases as a risk, but arguably a safer risk than to pursue the same endeavour in real life.

Related research indicates that the online and virtual world can often provide users with a sense of "online invincibility" (McCarthy, 2019), characterised by a change in attitudes toward risk perception and performing risky behaviour (similar to the online disinhibition effect; Suler, 2004). As such, one could expect the online or virtual world to provide users high in Emotionality to pursue forms of improvement that they would not in the offline world. Furthermore, in the context of gaming specifically, research also highlights that conditions of online anonymity can often boost gamer performance (Hénaff et al., 2015), therefore to a user high in Emotionality, the protective property of the online world (combined with anonymity) could support their preference for design features which lead to player development and enhancement.

In the case of high Extraversion predicting preference for Improvement design features, studies have indicated that Extraversion is correlated to strong motivations to learn and develop (Major et al., 2006). The tendency to learn and develop would complement the preference for design features which enable users to enhance or develop their performance. For example, the Item Power-up design feature serves the function of assisting users in acquiring enhancements that can improve their performance or character abilities. In the case of user with high Extraversion, this design feature would complement their preference and trait-disposition to wanting to improve and develop. Moreover, as high Extraversion is associated with increased self-awareness (Ben-Artzi & Hambuger, 2001), Extraverts would be expected to be interested in their own development, via a recognition of where improvements can be made, further complementing an increased preference for Improvement design features. A Progress Bar design feature, which visually illustrates to what extent progress has been made to achieving a milestone, would provide users high in Extraversion insight into how close they are to reaching a stage of improvement.

Extraverts are also expected to thrive in social surroundings wherein social interactions take place (Eid et al., 2003), which may explain why the Trade design feature would be

highly preferable. For example, the Trade design feature possess an implicit/intrinsic social component (i.e., the necessary counterparty when buying or selling), which users might (unconsciously) factor into their preference rating. In the context of high extraversion, it is reasonable to also expect that a design feature which is fundamentally operational based on social interaction would also be of increased preference.

5.7.2.6 Personality and Accessibility dimension preference

Preference for Improvement design features (Table 17) was predicted by the personality dimensions of Honesty-Humility, Emotionality, Extraversion and Agreeableness (Table 14), with higher scores of Agreeableness, Emotionality, and Extraversion correlating with increased preference for Accessibility design features, while higher scores of Honesty-Humility corresponding to decreased preference for Accessibility design features. As described previously, users who score high in Agreeableness are likely to be more accepting and cooperative (Ferguson et al., 2020), while users who score high in Extraversion are highly sociable and often pursue positions of leadership (Campbell et al., 2003). Given these trait dispositions, users may seek out and prefer Accessibility design features, given the functions they serve at providing guidance and likely detail on how to make use of design features that are of more dispositional interest. For example, how to perform cooperative functions such as reviving a teammate (in the case for high Agreeableness), or how to become a team captain and issue commands (in the case for high Extraversion). Similarly, in the case of high Emotionality, which is characterised as an increased sensitivity to life stress and anxiousness to danger (Ashton & Lee, 2004), Accessibility design features would provide users with the information on how to best prepare for adversity within the virtual world, and in this sense a user with high Emotionality would find Accessibility design features, particularly Walkthrough and Beginners luck, highly preferable.

Of the four personality dimensions which predicted preference for Accessibility design features, the Honesty-Humility dimension was negatively correlated, such that higher levels corresponded to decreased preference. According to Ashton and Lee (2004), high scores of Honest-Humility would be characterised by rule adherence, and the absence of the manipulation of others and entitlement. Thus, one would expect that a user who scores high in this personality dimension to report an increased preference for

Accessibility design features, given that in addition to providing familiarity with the virtual world it is often the case that game rules are also conveyed to a user.

The results could be explained by genre-specific experiences of Accessibility features, wherein behaviours which a high Honesty-Humility user may be averse to performing (Weller & Thulin, 2012). For example, as these design features serve the function of informing a user on how to perform tasks and operate effectively within the game, if to operate effectively a user is guided on how to conceal or manipulate (Holden, Zeigler-Hill, Pham, and Shackelford, 2013), how to deceive (Gylfason et al., 2016), how to accumulate or steal (Lee et al., 2005), or how to mock, taunt or gloat (Torres-Marín et al., 2019), users may be averse to such features given their trait tendencies to avoid antisocial behaviour (Ashton & Lee, 2004). In the present study, of the 14 gaming Reddits participants were recruited from, six were generic communities, while five were games heavily oriented towards competition, and defeating opponents (GoW, Rainbow6, Red Dead, RocketLeague, and Witcher). To illustrate, all games (except for Rocket League) have received content ratings of PEGI 18+, which is an adult classification applied "*when the level of violence reaches a stage where it becomes a depiction of gross violence, apparently motiveless killing, or violence towards defenceless characters"* (Robertson, 2014). It could therefore be argued that users who reported high Honesty-Humility but decreased preference did so due to their experience with Accessibility design features not consisting with their trait dispositions (data to this effect was not collected).

5.7.3 Study improvement suggestions

One area of improvement for the present study concerned the addition of 10 new design features to the DFPS (Table 6). While the original 37 design feature vignettes were assessed for representativeness in the previous study, the new additions were not. It could be argued therefore that the new additions may not possess the degree of representativeness as the original vignettes, thus might not be suitably included in the DFPS or applied to further research. However, the findings from the previous study highlighted several considerations that, if followed, are believed to provide adequate confidence in the representativeness of the new additions. For instance, the new additions were assessed for readability, and achieved similar scores to the original vignettes. Furthermore, the number of words each new addition consisted of was similar

to that of the originals, and were worded in the possessive form, thereby emulating the wording strategies of the originals.

Table 18
New additions to DFPS

Design feature	**Vignette**
Punish	Being able to punish an opponent once you have defeated them
Friend Invite	Having the option to invite friends to play with you
Communal discovery	Being able to complete tasks and learn new things as a community
Bonus	The scenario in which an unexpected or additional reward is received
Lottery	The scenario in which you receive a reward by luck
Scarlett letter	The element where other players are made aware of when you are stuck/failing
Performance graphs	Being able to view your performance level in a graph
Boss Battles	Overcoming an enemy/task at the end of a level that has marginally increased in difficulty
Beginners luck	Help in achieving a high rate of success when completing the first few tasks
Vote	Having the opportunity to vote on something (e.g., map, weapon, rules)

The sample might also have been less skewed to one gaming community. In the previous study, a suggested improvement was to diversify the participant sample to improve generalisability and control for community-specific perspectives, attitudes, or perceptions. The present study increased the number of online gaming communities from which participants were recruited to a total of 14 Reddits. Despite this, a significant portion of participants were recruited from the WoW Reddit (Table 15). Given the previously discussed concerns with focusing on one gamer community, future research will benefit from not only expanding the number of communities from which players are recruited, but also controlling for the proportion of participants from each community.

5.7.4 Summary

The primary outcome of the present study confirmed how user motivation and personality would predict preference for design features. Building on these findings, future research would benefit from further expanding the dimensions of individual differences measured in relation to user design feature preference, such that the adaptation of Gamification design can be more extensively suited to the needs of the user. One example could be to focus on the individual difference of Gamefulness, to gauge which aspects of a game experience a user finds valuable. Defined as the degree to which a user experience emulates that which would be experienced when playing games. (Deterding et al., 2011), Gamefulness could provide a more contextually suitable guide on how Gamification can be adapted to increase user engagement, given its conceptual origin stems directly from Gamification (Högberg et al., 2019). The present study also generated a nine-factor structure to the DFPS, following an EFA. Keeping in line with appropriate model development practices, an important next step would be to subject the DFPS to a CFA, such that the underlying factor structure can be further tested and validated.

6. Study Three

6.1 Abstract

User motivation and user personality were both found to possess a significant influence on user design feature preference in the previous study. One criticism of employing these areas of individual differences as predictors of design feature preference is that they do not directly originate from the field of Gamification (Högberg et al., 2019). Subsequently, adapting insights from personality or motivation to the context of Gamification, may result in a loss of nuance that could be critical to understanding how to maximise user engagement. The concept of Gamefulness provides a solution to this issue, being defined as the degree to which a user experience emulates that which would be experienced when playing games. Using an online cross-sectional design, a total of 1111 completed the DFPS along with the Gameful Experience Questionnaire (Högberg et al., 2019) to examine how Gamefulness could influence user design feature preference. The DFPS was subject to a CFA, which repeated a nine-factor solution, although some design features were re-allocated to new factors following the EFA conducted in the previous study. Specifically, three design features were added to the Improvement factor, two design features were added to the Competition factor, and one design feature was added to the Cooperation factor. Multiple regressions highlighted a significant impact of all Gamefulness dimensions on user design feature preference. Implications and further research are discussed, such as how the relationship between individual differences and design feature preference should be tested using more objective measurements of user engagement.

Introduction

The present study aimed to further develop the DFPS by assessing the extent to which it relates to the individual difference of Gamefulness.

6.2.1 Gamefulness

Unlike motivation and personality, both of which are areas that take from broader psychology and are used to interpret the behaviours of players, Gamefulness is a concept which stems directly from the area of Gamification (Högberg et al., 2019). The concept is defined as the degree to which a user experience emulates that which would be experienced when playing games (Deterding et al., 2011). For example, aspects such as such as challenge, difficulty, clarity of purpose (Huotari & Hamari, 2012; Deterding, 2015), engagement in artificial competition (Salen and Zimmerman, 2004), and objective fulfilment (Juul, 2005). As such, Gamefulness is a metric which provides researchers with insight to what aspects of a gaming experience a user values. In this sense, if an experience is more Gameful from the perspective of the user, then it is expected to also be more engaging, and if an experience is less Gameful from the perspective of the user, then it is likely to be less engaging.

In the context of Gamification design, Gamefulness offers a similar type of insight that can be gained from understanding user motivation and personality, such that individual characteristics can be leveraged to inform how Gamification could be adapted and tailored to the needs of the end user. For example, in understanding how design feature preference relates to user motivation, Gamification could be designed to more effectively fulfil the motivations of that user. Similarly, in understanding how design feature preference relates to user personality, Gamification could be designed to more effectively appeal to the trait dispositions of that user. In the context of Gamefulness, by understanding how design feature preference relates to the user's perception of what constitutes a Gameful experience, Gamification could be designed to more effectively curate a Gameful experience that appeals to that user.

Though similar in how they could be studied, compared with motivation and personality, Gamefulness could provide a more appropriate metric for predicting user preference in the context of adapting Gamification design. For example, most research used to inform how Gamification can be adapted tends to rely on studies that focus on gaming which

takes place on a specific platform, such as PC or mobile gaming. Hamari and Koivisto (2014) argue that Gamification cannot emulate the immersive sensory experience provided by platform gaming (due to technological and contextual limitations), therefore research which focuses on player behaviour observed in the virtual world may not be suitable to infer how players may behave in the real world within gamified contexts (Högberg et al., 2019). Recognising the differences in user behaviour that emerge between platform gaming and gamified contexts, Högberg et al. (2019) developed the Gameful Experience Questionnaire, which is an instrument design to measure what aspects of gaming are most valued by a user. The model proposes seven fundamental dimensions of Gamefulness (Table 19), each of which indicates which aspects of system interaction constitute a Gameful experience from a user perspective. Understanding what elements of a game experience the user values most could inform how Gamification design can be tailored and adapted to curate a more gameful experience, leading to an increase in overall user engagement.

Table 19

Gameful Experience Questionnaire (Högberg et al., 2019)

Accomplishment	**Challenge**	**Competition**	**Guided**	**Immersion**	**Playfulness**	**Social-experience**
Experiencing demand or drive for successful performance, goal achievement, and progress	Experiencing demand for great effort to be successful	Experiencing rivalry towards one or more actors (self, other person, service, or group	Experiencing being guided on how, what, and when to perform or improve a target behaviour	Experience where all attention is taken over and one is absorbed in what they are doing	Experiencing involvement in voluntary and pleasurable behaviours free from spontaneously created rules	Experiencing the presence of users, and acting as or encountering service-created social actors

6.2.2 Suitability for Confirmatory Factor Analysis

Further validation of the DFPS model would also require the nine-factor solution generated in the previous study to be subject to a CFA to test whether factor loadings fit consistently with the previous model (Bédard et al., 2015). As part of the wider factor analysis statistical method, CFA is a statistical technique which seeks to verify the factor structure of a set of variables (Suhr, 2006). The primary benefit of performing CFA is to enable the researcher to test that a latent construct exists. In the context of the present study, CFA would serve as a method to confirm and verify the nine-factor solution generated in the previous study. The suitability of CFA for the present study could be

explained in four main ways. The first, is that CFA is less suited at the early scale development stage, as it does not indicate how well individual items (i.e., design feature vignettes) load onto each factor (Kelloway, 1995). Generally, the use of CFA in an exploratory capacity is inappropriate (Brown, 2003). However, if a factor structure already exists and the researcher already possesses a hypothesis about the factors that are to be confirmed, then it is highly suitable (Jöreskog, 1969). In the context of this PhD, the a priori hypothesis is the nine-factor solution generated in the previous study by the Principal Axis Factoring (Table 20).

Second, is that CFA is most suited for new sample that was not used to obtain the initial factor structure (Fokkema & Greiff, 2017; Willmer et al., 2019). As the initial factor structure was generated by PCA in the previous study, CFA would be performed in the present study with a new sample. The third, was that CFA is conducted on a sample size of at least n > 200. As per the first and second study of this PhD, achieving a sample of more than 200 participants was achievable, given the role of online recruitment and high participation rate when using Reddit as a recruitment platform. Finally, the existing model scored above .3 for all respective items, indicating that each variable (design feature) possessed sufficient common variation to retain in the factor solution (Högberg et al., 2019).

Table 20

EFA nine-factor solution (DFPS)

Difficulty	**Competition**	**Reward**	**Accessibility**	**Loss**	**Expression**	**Cooperation**	**Improvement**	**Narrative**
Behavioural momentum	Leaderboard (Feedback)	Badges	Tips / Hints	Restriction	Avatar	Shared Goal	Item Power-up	Storyline
Levels	Leaderboard (Competition)	Medal	Walkthrough	Depletion	Design / Editing / Customisation	Complementarity	Currency	Cut Scenes
Game Goal	Rank	Trophy	Beginners Luck	Demotion	Profile	Communal Discovery	Tokens	Decision Making
Game Objective	PVP		Notifications	Scarlett	Vote	Voice Chat	Trade	Bonus
Boss Battle	Points			Item Degradation	Emotes	Friend Invite	Progress Bar	
	Performance Graph			Barriers		Text Chat		
	Dashboard			Lottery				
				Punish				

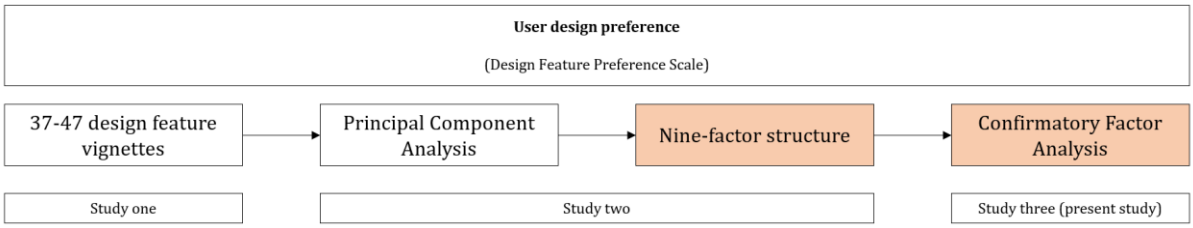
6.3 Rationale

The scope of the present study built upon the findings of Study one and Study two. In the first study, the DFPS was developed to measure user design feature preference. In the second study, the DFPS was validated and used to explain how design feature preference could be predicted by user motivation and personality. The present study sought to further validate the DFPS and assess how design feature preference could be predicted by Gamefulness. As Study two highlighted how preference for some design feature dimensions predicted variance in user engagement (Expression, Improvement, and Accessibility), the present study maintained this focus when assessing how effective Gamefulness could be in predicting design feature preference. A flow diagram illustrating the direction of this relationship can be found in Figure 17.

Figure 17 – Flow diagram illustrating direction of variable relationships

The present study sought to further validate the DFPS model by performing a CFA. A flow diagram illustrating how performing the CFA fits with the overarching aims of the PhD can be found in Figure 18.

Figure 18 – Flow diagram illustrating context of CFA within overarching research aims



6.4 Research aim

The present study aimed to assessed how the individual difference of Gamefulness could predict user preference for the EIA design features, given that in the previous study, these dimensions were found to significantly predict variance in user engagement. A secondary aim was to subject the existing DFPS model to a CFA.

6.5 Methods

6.5.1 Design

The present study employed a within-subjects cross-sectional online survey methodology to identify how motivation and personality relate to user design preference. Participants were asked to complete a user engagement scale, a Gamefulness scale and provide preference ratings for a series of design feature vignettes. As such there were three variables of interest: reported user engagement as measured by play duration and play frequency questions: responses to the Gameful Experience Questionnaire (Högberg et al., 2019) which measured user Gamefulness; and responses to the DFPS, which measured user preference for 47 design features.

6.5.2 Participants

Players were recruited from various gaming Reddits (see Table 21). A total of 1111 players participated, with 89% male, 9% female, and 2% transgender or not identifying. Ages ranged from 16 to 69 (Mean age = 25.55 years; SD = 5.32). The average level of gaming experience for the sample was 15 years (SD = 6.98). 42% of participants were employed, 37% were students, 13% were actively seeking employment, and 8% were unemployed. Of the total sample, 46% were from North America, 31% from Europe, 12% from Asia, 6% from South America, and the remaining from Oceania or Africa. 76% of participants reported most use of a PC platform to play video games, followed by consoles (XBOX, Playstation, Nintendo) at 13%, and mobiles and tablets at 1%.

Table 21

Reddit sample distribution (largest to smallest)

Reddit	***(n)***
Steam	474
Skyrim	188
Borderlands	110
TrueGaming	103
AssassinsCreed	88
GamePhysics	77
OverWatch	24
GTAOnline	14

ClashRoyale	8
No Reddit (via friend referral)	8
PCMR	7
RocketLeague	7
GameDeals	1
Doom	1

6.5.2.1 Eligibility criteria

Participants were only eligible to participate if they met two main criteria. The first, was that participants must have regularly played video games for at least two hours per week, which in previous gaming research has been regarded as the minimum time spent playing to qualify as being a gamer (Kolo & Braun, 2004). It was expected that players possess the prerequisite experience to understand the functional representation of Design features, therefore ratings of representativeness would be more reliable than those given by non-players. The second criteria were for participants to have been aged 18 years or older. As the study exclusively employed online methodologies, there was no way for researchers to obtain parental consent for underage participants, therefore only those of the legal age to consent were able to participate.

6.5.2.2 Participation incentive

All participants were automatically enrolled into a free game giveaway, wherein the successful winner would be awarded a game of their choice, limited to $77 or £55 and purchasable only from a reputable online seller (such as STEAM, Origin, or the XBOX marketplace).

6.5.3 Materials

6.5.3.1 Design feature preference

Participants were required to complete the DFPS, which required participants to report to what extent they found a given design feature fun, motivating, useful, and preferable (Lopez and Tucker, 2019). A total of 47 design features were functionally represented via vignettes, each of which had been generated from a literature review (Arnab et al., 2015; da Rocha Seixa et al., 2016; Hamari & Lehdonvirta, 2010; Lameras, 2017; Marczewski,

2015; Nacke, 2018; Orji et al., 2018; Rocha et al., 2008; Sailer et al., 2017; Tondello et al., 2017; Werbach & Hunter, 2012), and previous study validation (see Table 16). Each vignette was limited to 17 words and achieved an average Flesch reading ease score of 53.31 and a Flesch-Kincaid grade level test score of 8.85, suggesting that the vignettes could be easily read by 13-14-year olds. Together, all design features achieved an alpha of (a = .934) suggesting high internal consistency. Ratings were made via a 5-point Likert scale (1 – strongly unfavourable, 5 – strongly favourable) that was designed to emulate a conventional star rating system to induce participants to consider their ratings in a more meaningful capacity. Example vignettes include examples include the Complementarity design feature (being required to complete an objective or task only with the help and abilities of another player), and the Demotion design feature (being demoted and having your rank reduced after failing in some way).

6.5.3.2 Gamefulness

Participants were required to complete the Gameful Experience Questionnaire (Högberg et al., 2019), which measures to what extent different aspects of Gamefulness are valued by participants across 56 statements. Value ratings range from 1 – strongly disagree, 3 – somewhat disagree, 5 – somewhat agree, 7 – strongly agree. The scale measures seven dimensions of a Gameful experience; Accomplishment (experiencing demand or drive for successful performance, goal achievement, and progress), Challenge (experiencing demand for great effort in order to be successful), Competition (experiencing rivalry towards one or more actors), Guided (experiencing being guided on how (including what and when) to do, and on how to improve the target behaviour), Immersion (where all attention is taken over and one is absorbed in what they are doing), Playfulness (experiencing involvement in voluntary and pleasurable behaviours free from spontaneously created rules), and Social-Experience (experiencing the direct or indirect presence of people (real world or virtual), service-created social actors, and service as a social actor). Example statements include "Makes me feel immersed" and "Makes me feel like I'm developing something".

6.5.3.3 User engagement

User engagement was measured by play duration (how long a typical play session lasts) and play frequency (how often a participant plays in a week), as has been used in previous gaming research (Brunborg et al., 2014).

6.5.3.4 Demographics

Participants were asked to submit standard demographic data, such as age, ethnicity, gender, education status, country of domicile, employment status and marital status. In addition, general gaming related demographics will also be collected, such as favourite game title, favourite game genre to play, most used gaming platform, and for how long they have played video games (in years).

6.5.4 Procedure

Participants access the survey via a URL link provided in the recruitment advertisement, after which they were presented with the study information sheet. After providing consent and completing the demographics sheet, participants were required to report their level of engagement, followed by the completion of the Gameful Experience Questionnaire. Thereafter, participants completed the design feature preference rating task, and were finally debriefed.

6.5.5 Ethics approval

Ethical approval was granted by Birmingham City University's research ethics committee under the reference code Lally/2024/Am/2019/Aug/BLSS.

6.6 Results

There were two analyses conducted for the present study. The first, was to subject the nine-factor solution to a CFA. The second, was to assess how Gamefulness would predict preference for the design feature dimensions of Expression, Improvement, and Accessibility. Notably, these design feature dimensions were selected as per the outcomes of the previous study, which identified that variance in user preference for these design feature dimensions significantly determined levels of user engagement. Descriptive statistics showing mean scores mean scores for user engagement, user valued Gamefulness, and design feature preference, as well as indications of normality (Skewness and Kurtosis), and the number of participants that provided responses is outlined in Table 35.

Table 35

Distribution statistics and mean scores for user engagement, Gamefulness, and design feature preference

Variable	Mean	SD	Skew (SE)	Kurt (SE)	*n*
User engagement					
Play duration	21.77	15.59	1.82 (.073)	4.47 (.147)	1111
Play frequency	7.65	5.29	4.61 (.073)	35.25 (.147)	1111
Gamefulness (Gameful Experience Questionnaire)					
Accomplishment	44.33	7.25	-0.89 (.073)	1.51 (.147)	1111
Challenge	41.95	8.29	-0.69 (.073)	0.81 (.147)	1111
Competition	32.43	8.85	-0.35 (.073)	-0.21 (.147)	1111
Guided	29.43	7.27	0.10 (.073)	-0.11 (.147)	1111
Immersion	48.46	8.63	-0.68 (.073)	0.50 (.147)	1111
Playfulness	53.48	7.09	-1.08 (.073)	1.86 (.147)	1111
Social Experience	36.41	10.69	-0.36 (.073)	-0.24 (.147)	1111
Design features					
PVP	12.38	4.84	-0.40 (.073)	-0.44 (.147)	1111
Leaderboard (Competition)	13.45	4.92	-0.45 (.073)	-0.68 (.147)	1111
Punish	7.44	5.00	0.37 (.073)	-0.93 (.147)	1111
Complementarity	10.44	5.32	0.10 (.073)	-1.16 (.147)	1111
Shared Goal	13.98	4.67	-0.68 (.073)	-0.30 (.147)	1111
Trade	14.43	4.49	-0.57 (.073)	-0.48 (.147)	1111
Friend Invite	17.13	3.96	-1.86 (.073)	3.82 (.147)	1111
Text Chat	14.67	4.14	-0.72 (.073)	0.11 (.147)	1111
Voice Chat	13.74	4.82	-0.66 (.073)	-0.28 (.147)	1111
Emotes	10.53	5.18	-0.04 (.073)	-0.97 (.147)	1111
Communal discovery	14.18	4.78	-0.85 (.073)	0.02 (.147)	1111
Trophy	13.82	4.88	-0.75 (.073)	-0.02 (.147)	1111

Badges	12.65	5.20	-0.53 (.073)	-0.50 (.147)	1111
Medal	12.29	5.32	-0.45 (.073)	-0.66 (.147)	1111
Tokens	14.08	5.22	-0.82 (.073)	-0.24 (.147)	1111
Item Power-up	14.77	4.84	-0.88 (.073)	0.26 (.147)	1111
Bonus	14.23	4.54	-0.82 (.073)	0.25 (.147)	1111
Lottery	6.07	4.32	0.88 (.073)	0.09 (.147)	1111
Depletion	8.02	4.76	0.21 (.073)	-1.02 (.147)	1111
Restriction	5.50	4.25	1.16 (.073)	0.54 (.147)	1111
Demotion	8.28	4.90	0.31 (.073)	-0.88 (.147)	1111
Points	14.86	4.16	-0.79 (.073)	0.49 (.147)	1111
Progress Bar	15.34	4.04	-1.01 (.073)	0.75 (.147)	1111
Leaderboard (Feedback)	13.71	5.09	-0.47 (.073)	-0.70 (.147)	1111
Scarlett letter	8.01	5.03	0.28 (.073)	-1.05 (.147)	1111
Performance graphs	13.73	5.01	-0.61 (.073)	-0.34 (.147)	1111
Walkthrough	10.08	4.42	0.13 (.073)	-0.71 (.147)	1111
Tips / Hints	11.23	4.30	-0.21 (.073)	-0.48 (.147)	1111
Notification / Prompts	9.36	4.60	0.07 (.073)	-0.70 (.147)	1111
Cut Scenes	14.76	4.62	-0.80 (.073)	-0.08 (.147)	1111
Storyline	16.35	4.13	-1.59 (.073)	2.51 (.147)	1111
Currency	15.27	4.54	-0.97 (.073)	0.30 (.147)	1111
Item Degradation	6.87	4.50	0.65 (.073)	-0.57 (.147)	1111
Dashboard	14.78	4.37	-0.94 (.073)	0.59 (.147)	1111
Behavioural Momentum	16.23	4.10	-1.40 (.073)	2.00 (.147)	1111
Levels	12.26	5.23	-0.66 (.073)	-0.11 (.147)	1111
Barriers / Access	14.08	5.07	0.26 (.073)	-0.86 (.147)	1111
Game Goal	13.80	4.55	-0.72 (.073)	0.06 (.147)	1111
Game Objective	7.96	4.94	-1.05 (.073)	1.04 (.147)	1111
Boss Battles	14.51	4.01	-1.49 (.073)	1.99 (.147)	1111
Beginners luck	14.08	4.79	-0.01 (.073)	-0.89 (.147)	1111
Design / Editing / Customisation	15.83	3.91	-1.14 (.073)	0.63 (.147)	1111
Decision Making	16.77	3.77	-1.60 (.073)	2.42 (.147)	1111
Vote	8.86	4.88	-0.66 (.073)	-0.35 (.147)	1111
Avatar	15.31	4.68	-0.83 (.073)	-0.29 (.147)	1111
Profile	16.16	4.25	-0.49 (.073)	-0.74 (.147)	1111
Rank / Status	12.61	4.84	-0.74 (.073)	-0.14 (.147)	1111

Notes: SD = Standard deviation, Skew = Skewness, Kurt = Kurtosis, SE = Standard error, n = Sample size.

6.6.1 Analysis One – Confirmatory Factor Analysis

Using a full a priori specified model (Table 20), a CFA was conducted on the nine-dimension solution using Jamovi (version 2.0.0.0) with a maximum likelihood estimation. The hypothesised measurement model had fit indices χ^2 (n=1111) = 5319, p < .001, χ^2/df = 998, CFI = .790 (comparative fit index), RMSEA = .062 (90% confidence interval of .060 to .064; root mean square error of approximation), and SRMR = .075 (standardized root mean squared). This model possessed a moderately good fit, with the SRMR meeting the recommended cut off value of .08 (Hu & Bentler, 1999) and RMSEA meeting the cut off value of below 0.07 (Steiger, 2007). However, the CFI fell short of the .90 recommended cut off value (Hu & Bentler, 1999).

Table 20

EFA nine-factor solution (DFPS)

Difficulty	**Competition**	**Reward**	**Accessibility**	**Loss**	**Expression**	**Cooperation**	**Improvement**	**Narrative**
Behavioural momentum	Leaderboard (Feedback)	Badges	Tips / Hints	Restriction	Avatar	Shared Goal	Item Power-up	Storyline
Levels	Leaderboard (Competition)	Medal	Walkthrough	Depletion	Design / Editing / Customisation	Complementarity	Currency	Cut Scenes
Game Goal	Rank	Trophy	Beginners Luck	Demotion	Profile	Communal Discovery	Tokens	Decision Making
Game Objective	PVP		Notifications	Scarlett	Vote	Voice Chat	Trade	Bonus
Boss Battle	Points			Item Degradation	Emotes	Friend Invite	Progress Bar	
	Performance Graph			Barriers		Text Chat		
	Dashboard			Lottery				
				Punish				

In response to this, modification indices were consulted to improve on the model fit. Of these indices, higher scores were consulted and appropriately modified, such that some design features were reallocated to different loadings due to their high modification scores (Table 22).

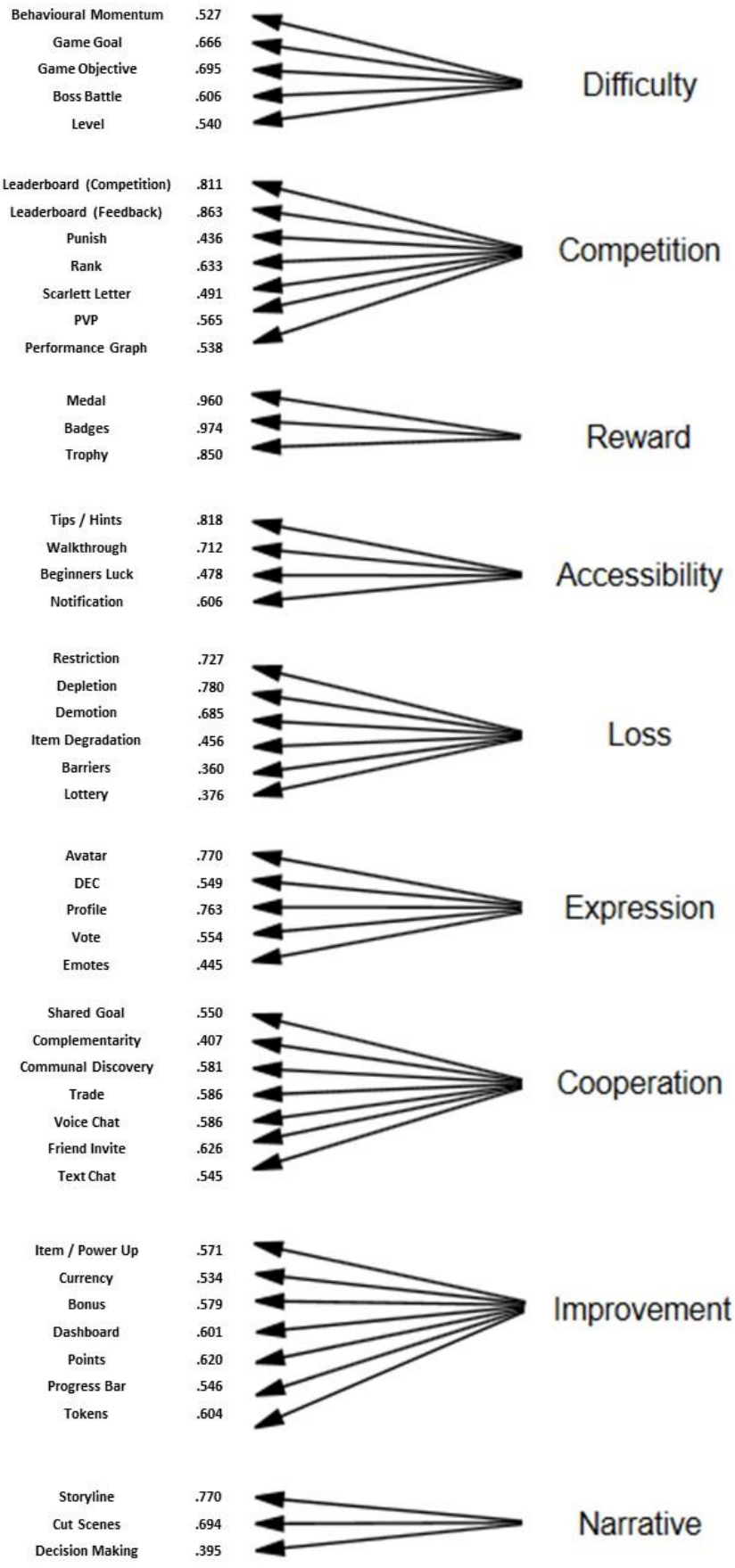
Table 22

Modification indices and changes made to CFA model

Design feature	Modification score	Original design feature dimension	New design feature dimension	Prior fit		Post reallocation	
				CFI	RMSEA	CFI	RMSEA
Dashboard	134.8993	Competition	Narrative	0.790	0.0624	0.792	0.0621
Points	118.1696	Competition	Improvement	0.792	0.0621	0.797	0.0614
Dashboard	106.8442	Competition	Improvement	0.797	0.0614	0.801	0.0609
Bonus	102.0316	Narrative	Improvement	0.801	0.0609	0.805	0.0609
Scarlett	100.51802	Loss	Competition	0.805	0.0609	0.806	0.0607
Trade	93.79321	Improvement	Cooperation	0.806	0.0607	0.807	0.0598
Punish	97.85198	Loss	Competition	0.807	0.0598	0.811	0.0593

Following this, the revised model had fit indices χ^2 (n=1111) = 4893, $p < .001$, χ^2/df = 998, CFI = .811, RMSEA = .059 (90% confidence interval of .057 to .060), and SRMR = .066. This model showed that all factor loadings were statistically significant, and possessed a better fit, however, still did not meet the CFI cut off considered a good fit. Nonetheless, all design features were included as they exceeded the inclusion threshold of .30 loading scores (Comrey & Lee, 1992). Additionally, the design feature loadings made theoretical sense and could still be justified following the model revision. The finalised DFPS model can be found in Figure 19.

Figure 19 – Finalised DFPS model


Behavioural Momentum .527
Game Goal .666
Game Objective .695
Boss Battle .606
Level .540
Difficulty
Leaderboard (Competition) .811
Leaderboard (Feedback) .863
Punish .436
Rank .633
Scarlett Letter .491
PVP .565
Performance Graph .538
Competition
Medal .960
Badges .974
Trophy .850
Reward
Tips / Hints .818
Walkthrough .712
Beginners Luck .478
Notification .606
Accessibility
Restriction .727
Depletion .780
Demotion .685
Item Degradation .456
Barriers .360
Lottery .376
Loss
Avatar .770
DEC .549
Profile .763
Vote .554
Emotes .445
Expression
Shared Goal .550
Complementarity .407
Communal Discovery .581
Trade .586
Voice Chat .586
Friend Invite .626
Text Chat .545
Cooperation
Item / Power Up .571
Currency .534
Bonus .579
Dashboard .601
Points .620
Progress Bar .546
Tokens .604
Improvement
Storyline .770
Cut Scenes .694
Decision Making .395
Narrative

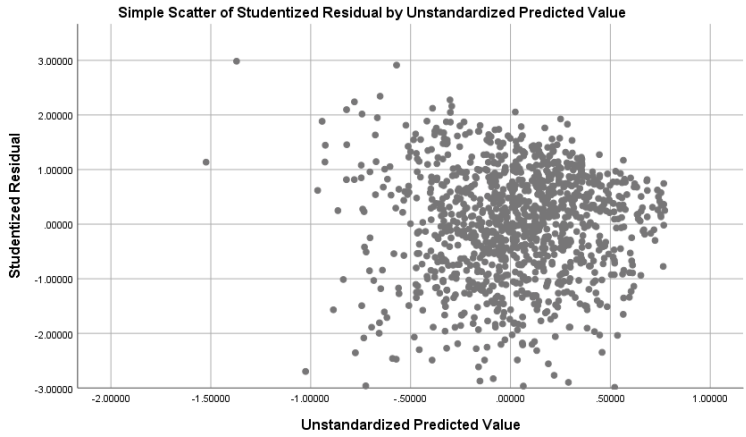
6.6.2 Analysis Two – User Gamefulness and design feature preference

6.6.2.1 Expression design features

Following the revised model, a multiple regression was conducted to assess how Gamefulness dimensions from the Gameful Experience Questionnaire predicted variance in Expression design feature preference. There was linearity as assessed by partial regression plots and a plot of studentized residuals against predicted values. There was independence of residuals as assessed by a Durbin-Watson statistic of 2.022. There was homoscedasticity as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values (Figure 20). There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1 (Hair et al., 2014), indicating that the independent variables (Gamefulness dimensions) were not highly correlated with each other. Measures to detect outliers revealed no studentized deleted residuals greater than ±3 standard deviations (Everitt & Skrondal, 2010), no leverage values greater than 0.2 (Huber, 1981), or values (influential points) for Cook's distance above 1 (Cook, 1977), indicating no outliers were present in the data.

The assumption of normality was met, as assessed by visual inspection of a Q-Q Plot (Appendix N). The multiple regression model significantly predicted user design preference, $F(7, 1094) = 21.420$, $P < .001$, adj. $R^2 = .12$. Preference for Expression design features was significantly predicted by the Gamefulness dimensions of Accomplishment ($B = .013$, $p < .050$), Competition, ($B = .010$, $p < .050$), Playfulness ($B = .018$, $p < .001$), and Social-experience ($B = .019$, $p < .001$). Specifically, reporting greater value for the Accomplishment, Competition, Playfulness, and Social-Experience dimensions of Gamefulness collectively correlated with increases in preference for Expression design features. The nonsignificant associations between the remaining three Gamefulness dimensions and Expression design features can be found in Appendix AAA4).

Figure 20 – Plot of Expression design feature preference and Gamefulness studentized residuals



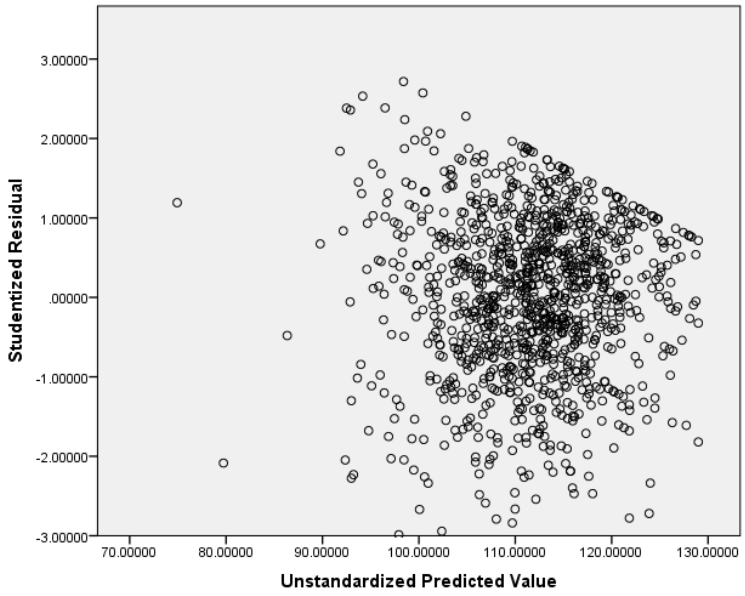
6.6.2.2 Improvement design features

A multiple regression was conducted to assess how Gamefulness dimensions from the Gameful Experience Questionnaire predicted variance in Improvement design feature preference. There was linearity as assessed by partial regression plots and a plot of studentized residuals against predicted values. There was independence of residuals as assessed by a Durbin-Watson statistic of 2.004. There was homoscedasticity as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values (Figure 21). There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1 (Hair et al., 2014), indicating that the independent variables (Gamefulness dimensions) were not highly correlated with each other. Measures to detect outliers revealed no studentized deleted residuals greater than ±3 standard deviations (Everitt & Skrondal, 2010), no leverage values greater than 0.2 (Huber, 1981), or values (influential points) for Cook's distance above 1 (Cook, 1977), indicating no outliers were present in the data.

The assumption of normality was met, as assessed by visual inspection of a Q-Q Plot (Appendix O). The multiple regression model significantly predicted user design preference, F (7, 1093) = 35.496, P < .001, adj. R^2 = .18. Preference for Improvement design features was significantly predicted by the Gamefulness dimensions of Accomplishment (B = .577, p < .001), Challenge (B = -.266, p < .050), Competition, (B = .339, p <.001), Playfulness (B = .188, p < .050), and Social-experience (B = .234, p < .001). Specifically, reporting greater value for the Accomplishment, Competition, Playfulness, and Social-experience dimensions of Gamefulness correlated with increases in

preference for Improvement design features, while greater value for the Challenge dimension of Gamefulness corresponded with a decreased preference for Improvement design features. The nonsignificant associations between the remaining two Gamefulness dimensions and Improvement design features can be found in Appendix AAA4).

Figure 21 – Plot of Improvement design feature preference and Gamefulness studentized residuals

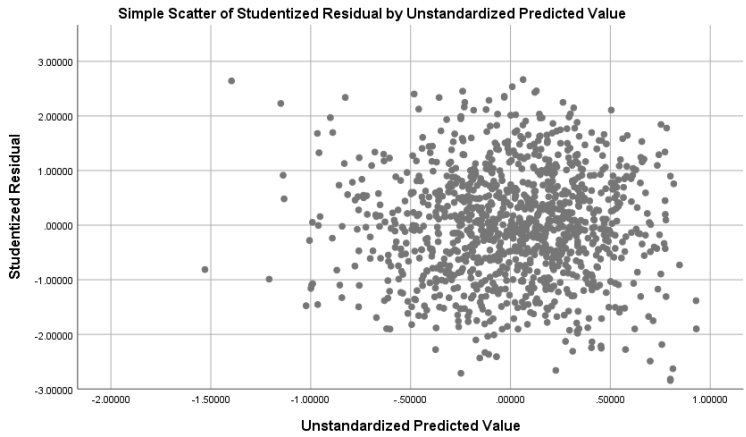


6.6.2.3 Accessibility design features

A multiple regression was conducted to assess how Gamefulness dimensions from the Gameful Experience Questionnaire predicted variance in Accessibility design feature preference. There was linearity as assessed by partial regression plots and a plot of studentized residuals against predicted values. There was independence of residuals as assessed by a Durbin-Watson statistic of 1.983. There was homoscedasticity as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values (Figure 22). There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1 (Hair et al., 2014), indicating that the independent variables (Gamefulness dimensions) were not highly correlated with each other. Measures to detect outliers revealed no studentized deleted residuals greater than ±3 standard deviations (Everitt & Skrondal, 2010), no leverage values greater than 0.2 (Huber, 1981), or values (influential points) for Cook's distance above 1 (Cook, 1977), indicating no outliers were present in the data.

The assumption of normality was met, as assessed by visual inspection of a Q-Q Plot (Appendix P). The multiple regression model significantly predicted user design preference, F (7, 1099) = 26.616, P < .001, adj. R^2 = .14. Preference for Accessibility design features was significantly predicted by the Gamefulness dimensions of Accomplishment (B = .016, p < .050), Challenge (B = -.029, p < .001), Guided (B = .049, p < .001), and Immersion (B = .007, p < .050). Specifically, reporting greater value for the Accomplishment, Guided, and Immersion dimensions of Gamefulness correlated with increases in preference for Accessibility design features, while greater value for the Challenge dimension correlated with decreased preference for Accessibility design features. The nonsignificant associations between the remaining three Gamefulness dimensions and Accessibility design features can be found in Appendix AAA4).

Figure 22 – Plot of Accessibility design feature preference and Gamefulness studentized residuals



6.7 Discussion

The primary aim of the present study was to understand how user preference could be predicted by the individual difference of Gamefulness (Table 19), which was argued as being a potentially more effective predictor than other individual differences (such as user motivation or personality). This aim was fulfilled, with the findings first indicating that variance in preference for EIA design feature dimensions were predicted the Gamefulness dimensions of Accomplishment, Challenge, Competition, Guided, Immersion, Playfulness, and Social- experience. Specifically, preference for Expression design features was predicted by variance in the Accomplishment, Competition, Playfulness, and Social-experience dimensions of Gamefulness; preference for Improvement design features was predicted by variance in the Accomplishment, Competition, Social-experience, and Challenge dimensions of Gamefulness; and preference for Accessibility design features was predicted by the Challenge, Accomplishment, Guided, and Immersion dimensions of Gamefulness. Instances where nonsignificant associations emerged can be found in Appendix AAA4.

The secondary aim was to further develop the DFPS model and confirm its factor structure. A CFA was performed to confirm nine-factor model structure. Following some revision of how six (of 47) design features loaded into broader dimensions, which are summarised in Table 23, the CFA revealed a moderately good fit. As such, the factor structure remained as: Difficulty (1) Competition (2) Reward (3) Accessibility (4) Loss (5) Expression (6) Cooperation (7) Improvement (8), and Narrative (9).

Table 19

Gameful Experience Questionnaire (Högberg et al., 2019)

Accomplishment	**Challenge**	**Competition**	**Guided**	**Immersion**	**Playfulness**	**Social-experience**
Experiencing demand or drive for successful performance, goal achievement, and progress	Experiencing demand for great effort to be successful	Experiencing rivalry towards one or more actors (self, other person, service, or group	Experiencing being guided on how, what, and when to perform or improve a target behaviour	Experience where all attention is taken over and one is absorbed in what they are doing	Experiencing involvement in voluntary and pleasurable behaviours free from spontaneously created rules	Experiencing the presence of users, and acting as or encountering service-created social actors

6.7.1 Confirmatory Factor Analysis

Following a CFA, the nine-factor model indicated a moderately good fit, however, modification indices identified how the fit could be improved by reallocating the loading of some design features. In total, six design features out of 47 were reallocated, with new additions to the dimensions of Competition, Cooperation, and Improvement (Table 23).

Table 23

Changes made to DFPS model

Design feature	**Description**	**Design feature dimension**	
		Before CFA	**After CFA**
Punish	Being able to punish an opponent once you have defeated them	Loss	Competition
Scarlett Letter	The element where other players are made aware of when you are stuck/failing	Loss	Competition
Trade	Being able to trade inventory items/currency with other players in exchange for items/currency	Improvement	Cooperation
Dashboard	Being able to access game information, such as your game history, resources, profile, friends list, achievements etc	Competition	Improvement
Points	Being able to see your progression in a number format	Competition	Improvement
Bonus	The scenario in which an unexpected or additional reward is received	Narrative	Improvement

6.7.1.1 Additions to the Competition dimension

Punish and Scarlet Letter

The Loss design feature dimension comprised of design features which serve a primary function of losing "something", encountered by a user following a failure of some sort. As such, both the Punish and Scarlet Letter design features loading onto the Loss dimension could be expected. For example, the Punish design feature, which enables users to punish other users following a failure of some sort, consists with the overarching characteristics of the Loss design feature dimension. Similarly, the Scarlet Letter design feature, which reveals when a user has encountered great difficulty and cannot surmount a given obstacle, could also be interpreted as a loss of some sort administered only after "failing" to complete a task. Following the CFA, modification indices suggested that both design

features would contribute to a better model fit if reallocated to the Improvement design feature dimension.

These changes were made as they are more appropriately suited to the Improvement design feature dimension. To illustrate, the Improvement design feature dimension comprises of design features which serve the function of enabling or curating competition between users. In the case of the Punish design feature, competition between users could be encouraged if there exists the option to punish rivals. Research indicates that nested within the motive to compete is the expression of forceful/aggressive behaviour, which could manifest in administering a punishment to the losing team (Franken and Brown, 1995). Similar assertions have been made in gaming motivation work, specifically the Killer motivation for Bartle's MUD player types (Bartle, 1966). At a more fundamental level, the Punish design feature requires the input of other users, such that punishment can only be administered by another user. In the absence of competition, it is difficult to justify the power provided to one user as being able to punish another without destabilising the balance of fairness that is necessary to keep users playing a game (Chen et al., 2020). Thus, conceptualising the Punish design feature as a Competition design feature consists with user psychology and the motive to compete, as well as the practical implementation of the design feature within a broader set of considerations (such as user engagement).

Similarly, the Scarlet Letter design feature is better suited to the Competition design feature dimension. In the initial dimension of Loss, the Scarlet Letter design feature was considered a punishment of some sort. However, in practice, one could argue that the revelation of player progress (or lack thereof) is only a punishment/loss if the user cares enough about that information being revealed. For instance, a user who plays with friends may not wish to have such information shared, as others may infer how skilful the user is (Hénaff et al., 2015). However, in the case of users who do not play with others, this becomes much less important. If the design feature is considered from the Competition perspective, its role at curating competition could be viewed as possessing a much stronger use case than its role as a form of loss (Butler et al., 2014). For example, competing teams would be provided real-time information on the progress of other players, with which a greater level of insight could inform their strategies. Additionally, teams could leverage the Scarlet Letter function, such that they could trick other players

into assuming they are facing difficulty as a form of deception (Rubin & Camm, 2013). In either case, it is argued that for the Scarlet Letter design feature as a form of Loss, it relies upon a broader context, which hints at competitive play. As such, it was also reallocated to the Competition design feature dimension.

6.7.1.2 Additions to the Cooperation dimension

Trade

The Improvement design feature dimension comprised of design features which serve a primary function of enabling user improvement of skills or abilities. In this sense, the Trade design feature, which enables the trading of items or in-game unlockables between users, would appropriately load onto the Improvement design feature dimension, given that trading items can (depending on what is traded) could lead to player improvement. Following the CFA, modification indices suggested that the Trade design feature would contribute to a better model fit if reallocated to the Cooperation design feature dimension. This change was made on the basis that the function of the Trade design feature is better suited to facilitating Cooperation than facilitating user Improvement. To illustrate, the function of Trade as facilitating user Improvement is largely dependent on what is being traded, therefore depending on the game in which the feature exists, items traded may not have any impact on user improvement. In contrast, if the Trade design feature is considered as a feature which facilitates cooperation between users (Witkowski & Kiba-Janiak, 2014), irrespective of what is traded, its characterisation as a Cooperation design feature persists beyond individual use cases.

6.7.1.3 Additions to the Improvement dimension

Dashboard and Points

The loading of the Dashboard and Points design features to the Improvement design feature dimension from the Competition design feature dimension was viewed as a more suitable placement when considering the primary function each design feature served. For example, though the Dashboard design feature provides users with a variety of in-game metrics (history, resources, friends list etc), what a user chooses to do with this information is unclear, therefore whether such information prompts a user to compete with others is uncertain. Similarly, the role of Points as facilitating competitive play is not

its primary function. In contrast, their functions align much more consistently with the Improvement design feature dimension, which comprises of design features that serve the function of enabling the improvement of user skill or ability. For example, the Dashboard design feature provides a variety of user metrics that inform the user on their level of progress/skill on a given metric (e.g., their resources or inventory) (Freitas et al., 2017). Similarly, the Points design feature provides users with a metric of their progress in a number format, further providing insight on how they can improve (or how close they are to improving/achieving a milestone) (Smiderle et al., 2020). Notably, the Points design feature closely corresponds to the Progress Bar design features, which also loads onto the Improvement design feature dimensions (the difference between the Points and Progress Bar being the way in which user progress is presented, with the former being progress in a number format, while the former being progress in a visual format). In both cases, the Dashboard and Points design features serve as forms of feedback, which can be used to inform user improvements in ability and skill.

Bonus

As with the Dashboard and Points design features, the Bonus design feature was also reallocated to the Improvement design feature dimension, again on the basis that doing so would be a more effective alignment of the design features function. For instance, the Bonus design feature serves the function of providing an unexpected or additional reward, however, the initial model loaded this design feature onto the Narrative dimension (which comprises of design features relating to the game narrative/story).

The receipt of an unexpected reward is not necessarily related to the game narrative/story, as this is mainly dependant on what is rewarded. Therefore, the loading of the Bonus to the Narrative design feature dimension was not entirely suitable, and only applicable in some instances (wherein the nature of Bonus is related to the game narrative). In comparison, an unexpected reward provided by the Bonus design feature can be more closely aligned with serving a function of user improvement, given that the reward will provide some form of enhancement for the user (though the type of reward that a user unexpectedly receives is largely dependent on broader game considerations).

6.7.2 User design preference and Gamefulness

The present study assessed how design feature preference could be predicted by dimensions of Gamefulness. In reference to the flow diagram illustrated earlier in this chapter, the focus on Gamefulness and design feature preference focused on the former part of Figure 17.

Figure 17 – Flow diagram illustrating direction of variable relationships

Table 24

Finalised loading of design feature dimensions that predicted variance in user engagement

Accessibility	Expression	Improvement
Tips / Hints	Avatar	Item Power-up
Being able to receive tips and hints when playing	Being able to represent yourself via a virtual model/sprite/signature	Being able to receive beneficial items / power ups after completing challenges or tasks
Walkthrough	Design / Editing / Customisation	Currency
Having the option to receive a step-by-step guide on how to complete tasks or play the game	The option to edit or design aspects of the game (e.g., avatar, environment, inventory)	Being able to spend your in-game money/currency on game content
Beginners Luck	Profile	Tokens
Help in achieving a high rate of success when completing the first few tasks	Being able to immediately convey several aspects of yourself to other players	Being able to earn tokens after completing challenges, that can be used to buy game content
Notifications	Vote	Progress Bar
Being able to receive key notifications and updates when playing	Having the opportunity to vote on something (e.g., map, weapon, rules)	Being able to see how close you are to reaching a milestone in a bar format
	Emotes	Dashboard
	Being able to express your emotion and feelings through your avatar behaviour (such as jumping or dancing)	Being able to access game information, such as your game history, resources, profile, friends list, achievements etc
		Points Being able to access game information, such as your game history, resources, profile, friends list, achievements etc
		Bonus The scenario in which an unexpected or additional reward is received

6.7.2.1 Gamefulness and Expression dimension preference

Preference for Expression design features (Table 24) was predicted by user value for the Gamefulness dimensions of Accomplishment, Competition, Playfulness, and Social-experience (Table 19), with increased value corresponding to increased preference. According to Högberg et al. (2019), users who value Competition seek rivalry with other users, which might be encouraged by the Avatar design feature. For example, as Avatars serve as a visual representation of a user that can be viewed by others, users are likely to compare their Avatar's appearance. Building on Social Comparison Theory (Festinger, 1954), which posits that there exists an internal desire to people to evaluate themselves against others, it could be argued that within the game space, such comparisons could be drawn between users and their Avatar appearance. For example, in addition to regular items a user has access to, game inventories can often include rare items, luxury items, or highly limited-edition items, which can only be acquired via performing completing highly difficult or time intensive challenges, or in-game payment. Studies indicate that the pursuit and acquisition of luxury or scarce items, such as designer clothing or expensive cars, can often be used as a competitive strategy to dominate or succeed against others (Hudders et al., 2016). Therefore, it is likely that such behaviour may emerge in virtual social context with respect to Avatar items, fuelling competitive dynamics between users.

Research indicates that social comparison can often drive user engagement, such that both upward and downward contrasts (comparing against others who are perceived as better, thereby feeling inadequate, or comparing with others perceived as worse, thereby feeling better; Lyubomirsky & Ross, 1997) can incline users to play online games for longer (Esteves et al., 2021). As the previous study found that preference Expression design features can impact overall user engagement, the role of social comparison could provide another explanation for these findings.

At the core of this perspective exists the social context. One could expect users who value the Gamefulness dimension of Social-experience to also prefer Expression design features, which can often be used in social contexts. For instance, users who value the Social-experience dimension seek opportunities wherein they encounter and engage with others. The Emotes design feature, which enables users to express their emotions via Avatar behaviour (e.g., jumping for joy to convey excitement, or clapping to convey you are impressed), which are likely to enhance the experience of social interactions with

other users. Moreover, the Profile design feature, which immediately conveys aspects of the user to others, might also appeal to users who value the Social-experience, as it may inform user decisions to interact with other social actors (Bergstrom, 2021).

The value for Playfulness also corresponds to the preference for Expression design features, in particular the Design / Editing / Customisation design feature. Defined as an experience characterised by spontaneity and free from rules or restrictions, the DEC design feature serves the function of providing users with the ability to edit and customise aspects of the game, such as the user's Avatar, and inventory, as well as the game environment. In seeking experiences characterised by a form of in-game "freedom" (Hui & See, 2015), a preference for Expression design features, which serve the function of enabling creativity, customisability, and a freedom to customise, would be consistent and expected.

Surprisingly, the analyses revealed that users who prefer Expression design features also reported value for the Accomplishment dimension of Gamefulness, which is characterised as experiences that are oriented towards achieving milestones or making progress. As Expression design features do not primarily serve this function (perhaps much better suited to design feature dimensions of Improvement or Competition), the relationship found might be better attributed to how Expression design features promote aspects of Accomplishment indirectly. For example, several studies indicate that by design, game reward systems (Cruz et al., 2017) and dynamic difficulty adjustments (Liu et al., 2009; Zohaib, 2018) are two areas of game design which promote users to continue achieving. The former incentivises users to pursue an achievement, while the latter carefully curates a relative balance of difficulty and competency. Relating more to why a user seeks to achieve, research suggests that reward types that enhance the user experience (such as Avatar skins) tend to encourage repeated play (Macey & Hamari, 2018). Thus, one could argue that the relationship between Expression design feature preference and a value for Accomplishment builds more upon the involvement of Expression-style rewards as orienting the desire to accomplish, and less so the direct function served by Expression design features.

Alternatively, this relationship might also be explained by the time at which expression-style rewards are acquired by users. Several games offer rewards, such as character skins and collectibles that are characterised by a high level of rarity and scarcity (i.e., not many

players own them). These rewards are often only acquired by players who possess significant skill or have progressed substantially in the game world. As these players are more likely to have accomplished, the pursuit of expression-style rewards might increase in tandem with the emphasis on accomplishment.

6.7.2.2 Gamefulness and Improvement preference

Preference for Improvement design features (Table 24) was predicted by user value for the Gamefulness dimensions of Accomplishment, Challenge, Competition, and Social-experience (Table 19). Specifically, increased value for Challenge corresponded to increased preference for Improvement Design features, whereas increased value for Accomplishment, Competition, and Social-experience corresponded to decreased preference for Improvement design features.

Improvement Design features collectively serve the function of providing a user with the facility to enhance or improve their in-game performance. For example, a Progress Bar informs a user on how close they are to reaching a milestone, while the Item Power-up and the Bonus design feature can augment user abilities. Given that the Gamefulness dimension of Challenge is characterised as seeking great demand or effort to be successful, the user's desire to encounter challenge consists with an increased preference for Improvement design features, as both concepts share the broader aim of improving user performance and abilities. The relationship between game difficulty/demand and improvement is also supported by research, with studies highlighting that the increasing demand can often lead to increases in user skill and performance (Rodriguez-Guerrero et al., 2017).

In contrast, the Gamefulness dimensions of Accomplishment, Competition, and Social-experience did not emulate this relationship, with increased value corresponding to decreased preference for Improvement design features. One perspective as to why these inverse relationships emerged could relate to how the core characteristic of each dimension does not relate to user skill or performance improvement. For example, the dimension of Social-experience focuses on the need to interact with others, the Accomplishment dimension focuses on encountering demanding difficulty, and the Competition dimension focuses on experiencing rivalry with other users. A similar viewpoint can be observed from gaming motivation research, which has often

distinguished between a motive to improve individual competency, and the motive to socialise, accomplish, or compete (Dmetrovics et al., 2011; Kahn et al., 2015; Lafreniere et al., 2012; Lee et al., 2012; Yee et al., 2012;).

6.7.2.3 Gamefulness and Accessibility preference

Preference for Accessibility design features (Table 24) was predicted by user value for the Gamefulness dimensions of Accomplishment, Challenge, Guided, and Immersion (Table 19). Specifically, increased value for Accomplishment, Guided, and Immersion corresponded to increased preference for Accessibility design features, whereas increased value for Challenge corresponded to decreased preference for Accessibility Design features.

As defined by Högberg et al. (2019) the Gamefulness dimension of Accomplishment is characterised by experiencing the demand or drive for successful performance, achievement, and progress. In the context of preferring Accessibility design features, the positive correlation found could be explained by the immediate feedback provided that fulfils accomplishment recognition. For example, the Walkthrough design feature, which serves the function of providing a new user with a step-by-step guide on how to complete tasks, could be regarded as an effective source from which the user's progress is recognised and acknowledged. In practice, a Walkthrough design feature immediately confirms with the user whether the instructions they have followed during the walkthrough are correct (as a measure to guide the user to the next step). Moreover, the Notification design feature, which serves the function of providing a user with updates and notifications, would also contribute to an overall sense of accomplishment and recognition, given that the user is provided with tailored information on how they are progressing within the virtual world.

Valuing the Gamefulness dimension of Guided also coincided with an increased preference for Accessibility design features, given that the Guided Gamefulness dimension is characterised by a need to receive guidance on how to operate within the virtual world (i.e., what to do, how to do it, and when to do it), and that Accessibility design features serve to make the game more accessible to a user. For example, a Walkthrough design feature specifically seeks to provide users with a guided practice-based tutorial on how to operate in and become familiar with the virtual world. Similarly,

the Tips / Hints design feature seeks to provide further support during gameplay, while the Beginners Luck design feature actively adjusts difficulty levels for the user, such that the user can become further accustomed to the virtual world. In this sense, Accessibility design features closely fulfil that what would be considered as experiencing the Guided Gamefulness dimension.

In the context of the Immersion, users who value this Gamefulness dimension seek a form of absorption characterised by a sense of dissociation from the real world and their attention being taken over (Preston, 2012). When interpreting the relationship between valuing Immersion and increased preference for Accessibility design features, discerning a direct association is difficult, given that Accessibility design features do not directly serve functions to curate an immersive environment. However, as Accessibility design features serve the function of enabling the user to operate effectively and quickly in the virtual world (by way of providing opportunities and guidance on how to improve and learn), one could argue that this relationship focuses on a fundamental part of the overall user immersion process. For instance, achieving immersion is noted as requiring an optimal balance of competency and difficulty (Sigailov-Lanfranchi, 2019). As such, to become immersed, a user must be equipped with skills to overcome game difficulties.

While demand can be provided automatically by games which employ dynamic difficulty adjustments (Liu et al., 2009; Zohaib, 2018), Accessibility design features provide an opportunity (and the information) for users to improve their ability at overcoming game difficulties. Thus, an increased preference for Accessibility Design features could have been reported by users who highly value the Gamefulness dimension of Immersion because receiving the information and experience provided by Accessibility Design features would assist the user in achieving an Immersive experience.

The opposite was found for users who highly valued the Gamefulness dimension of Challenge. Defined as experiencing the demand for great effort wherein a user's ability is tested, the results found that higher value corresponded to a decreased preference for Accessibility design features. A potential explanation for this could be that users who value Challenge could regard Accessibility design features as providing "too much" information and knowledge on how to overcome game difficulties, such that tasks are not as challenging. Research has highlighted that players may not always seek the most effective or easiest way at overcoming an in-game difficulty, and despite the goal-oriented

dynamics of most games, part of the fun is overcoming the challenge at hand (Potter, 2015). Therefore, users who desire challenge could report decreased preference for design features which actively reduce how difficult a challenge may be. A prime example of this would be a decreased preference for the Beginners Luck Design features. Given its primary function of reducing difficulty, a user who seeks challenging experiences is unlikely to prefer this design feature, as interaction with this design feature would lead to unchallenging experiences. Users who report an aversion to receiving help to complete tasks might possess higher levels of intrinsic motivation, which is the motive to pursue goal-directed activity because of the enjoyment it brings, and not the reward that may follow goal fulfilment (Falk et al., 1999). Indeed, research highlights that players can often report enjoyment at the possibility of being defeated, in that outcome uncertainty and challenge can often increase excitement and be mediated by intrinsic motivation (Abuhamdeh et al., 2015).

6.7.3 Study Improvement suggestions

Building on suggestions from the previous study, a key design concern was the disproportionate number of participants that could be recruited from only one gaming community. In the previous study, the participant sample was skewed towards the WoW Reddit. In the present study, there was a majority portion of participants from the Steam Reddit. Although there did exist a skew in the sample, a key distinction between the skew of the previous study sample and the present study sample, is the nature of each respective online community. The WoW Reddit focuses exclusively on the WoW franchise, whereas in comparison, the Steam Reddit does not focus on any given title or genre, rather it focuses on the Steam marketplace, within which a variety of games can be discussed. Therefore, the community focus on a given title or genre was not a feature of the Steam Reddit, therefore the associated concerns of participant perspectives or attitudes as being predominantly characteristic of the given community, was not expected.

6.7.4 Summary

The primary outcome of the present study confirmed how value for Gamefulness would predict design feature preference. Combined with the previous study outcome of how user motivation and personality would predict design feature preference, the present study contributes further to the existing number of user metrics with which Gamification design can be informed and effectively adapted. Both the previous and present study utilised self-report responses to measure user engagement. However, when considering the limitations associated with relying on participant recall, the findings may not correspond to actual player behaviour. A persisting limitation with self-report measures is that participants may not correctly recall and report previous behaviour accurately (Demetriou et al., 2015), For example, if users who report strong trait tendencies for extraversion are expected to demonstrate increased engagement when interacting with Improvement design features, then in practice such a relationship must persist when observing the players real game behaviour. Thus, a more objective and verifiable measurement of engagement is necessary to validate the findings of the previous and present study. A suggested future direction with this information would be to experimentally validate these findings and further develop the psychometric properties of the DFPS, by using more objective user engagement measures. The DFPS was also further validated via the CFA, which maintained the nine-factor structure (with some small revisions; see Table 23). Building on the outcome of both the EFA and CFA, the DFPS demonstrated sound psychometric model properties, thereby making it a viable instrument to measure user design feature preference.

7. Study Four

7.1 Abstract

Central to the assessment of whether Gamification is successful is the variable of user engagement. Existing methods commonly used to measure user engagement require upon users to self-report their time spent playing, however, these methods are limited in reliability, and whether they represent natural play behaviour and can be used to effectively predict prospective play behaviour. One solution to this methodological limitation would be to observe the users' natural play behaviour. Building on the results of the previous studies, the present study sought to validate the relationship between design feature preference and user engagement by using a task-performance measurement of user engagement, wherein users would demonstrate their level of engagement. Using an online cross-sectional within-subjects experimental design, a total of 72 participants completed the DFPS and played five online-browser games, during which their overall play time was measured. Multiple Regression analyses revealed that that none of the previously identified relationships between user design feature preference and user engagement were found to be significant. Methodological implications for why these results occurred, as well as future research steps, such as the use of engagement measures which capture more natural play behaviour in comparison to experimental play behaviour.

7.2 Introduction

The overarching purpose of Gamification is to drive engagement with non-game tasks via the application of game design features (Suh et al., 2017), which is why a fundamental interest of Gamification research is to understand in what ways user engagement can be maximised. Within the Gamification literature, there are a variety of metrics used to measure engagement, most of which relate to self-report or task-performance. For example, self-report measures include interviews (Lounis et al., 2013) or questionnaires (Suh et al., 2016), and rely on the user to reflect and report their play habits, such as how often they play or for how long they play. In contrast, task-performance measures rely on the performance of users with respect to a gamified task. For instance, level of vegetable consumption in a health-Gamification study (Jones et al., 2014), frequency of module engagement in a learning-Gamification study (Landers and Landers, 2014), and technician efficiency at removing software bugs in an occupation-Gamification study (Arai et al., 2014).

In comparing the effectiveness of both methods, self-report measures are likely to provide researchers with a high degree of operational practicality as they are inexpensive, and easy to adapt and distribute for data collection (Hunter, 2012). In contrast, task-performance methods which directly measure player behaviour might require greater levels of resources to measure user engagement, given the focus on real-time user behaviour (i.e., how the user completes the gamified task) and resource required to accurately measure behaviour over time. Though self-report measures provide an accessible solution to researchers in measuring engagement (i.e., how long do you play or how often do you play, as asked of participants in the previous studies), self-report data is likely to be less reliable. Self-report measures rely on a user's subjective evaluation of their behaviour (Lucas, 2018), which when considering variance in introspective ability may not be an accurate representation of what is being reported (Demetriou et al., 2015). Participants may also provide answers pertaining to their behaviour in a manner which they perceive to be more favourable (Devaux & Sassi, 2016), which in the context of gaming research could lead to users underreporting their overall play activity (Jeong et al., 2018).

In contrast, task-performance measures which rely on the real-time behaviour of a user could be considered a more objective account of the user's play activity. Conceptualised

as a direct observation of how a user interacts with a system, real-time task performance provides researchers a layer of objectivity in assessing how a user reacts to a given system, to what extent system features are preferred, and to what extent the system engages the user (Lopez & Tucker, 2019). Beyond mitigating the subjectivity of self-report, levels of engagement that are recorded for participants who engage in task-performance are more appropriate in the context of Gamification research, as it exposes the user to the native environment in which Gamification would occur (i.e., interacting with the design features which would provide a more objective demonstration of the users engagement with the design features), and would likely possess greater representativeness of what would occur in a "real" gamified context (Lopez & Tucker, 2019). Thus, the activity and level of engagement measured is more likely to correspond to what would naturally occur.

In observing task-performance, the role of individual differences at influencing user engagement may also be more accurately identified as prompting a participant to completes tasks may encourage natural tendencies to emerge, given the more interactive stimuli of performing a task (as opposed to competing a questionnaire) and the behavioural arousal which occurs as a result. For instance, Speer et al. (2015) demonstrate that personality can be assessed when an individual responds to a real-life situation, such that natural traits and dispositions are likely to manifest. Similarly, research has been conducted which also demonstrates the reliability of observing motivation states of participants when completing a series of tasks in a performance environment (Matthews et al., 2001). As task-performance seeks to emulate a more natural play experience, one could reasonably expect a user to be more engaged, and thus any emergence of individual differences (such as personality traits or motivation states) may provide a more accurate representation of both the authentic traits of the user, and how such traits may be related to system interaction and subsequent user engagement. Additionally, such observations can be used to supplement and validate self-report measures. As such, the role of task-performance measures can not only benefit the reliability of conclusions drawn from user interaction with a gamified system, but also substantiate perceived relationships between user engagement and associated individual differences

However, while real-time task performance does offer a more objective account of engagement levels, it is a resource intensive measure. In comparison to self-report measures, real-time task performance requires a greater degree of user monitoring over a potentially longer period of data collection. For example, in a study which focused on the gamification of learning, students were asked to complete online activities throughout their undergraduate degree, with overall performance being measured across numerous behavioural metrics (attendance, training course completion, participation, and literacy competence) and intervals (Poondej & Lerdpornkulrat, 2016).

Thus, there exists a choice for which measurement of user engagement researchers could use, with self-report measures possessing potentially greater research practicality but lower objectivity, while task-performance possessing greater objectivity but lower practicality. A solution to this problem would be a measurement of engagement that is optimised to possess both the convenience and practicality of self-report measures as well as the objective elements of real-time task-performance. One such example would be online games.

The use of online games could enable researchers to collect objective measurements of engagement using fewer resources than conventional study-specific tasks. To illustrate, on the internet there are many websites which allow users to play games online at no cost. Given the wide range of genres, researchers could select a given game based on the profile of design features that are present within that game, thereby assisting with more specific assessments of engagement. For example, in assessing how users engage with a PVP design feature, researchers could instruct users to play a game with PVP elements. In the context of understanding user preference for design features (a core focus in Adaptive Gamification), possessing a research measure that can, where necessary, be adapted to specifically assess levels of user engagement when interacting with specific configurations of design features, could provide more targeted insights into how Gamification design might be adapted.

More targeted measures of user engagement that measure user activity in a more natural environment (i.e., when playing a game as opposed to self-reporting) might potentially provide richer insights into how user engagement is influenced and can be improved by the design of Gamification (i.e., which design features users are more receptive to). Moreover, as engagement is conceptualised as how long a user spends playing, measuring

engagement in this form will also be more objective than self-report, given the participant is not required to estimate or recall their playtime, and instead can be directly recorded. In this sense, the usage of online games in measuring engagement provides real-time demonstration of user activity, as opposed to a retrospective recall of user activity.

7.3 Rationale

The proposition of Adaptive Gamification is that user engagement is directly influenced by design feature preference, and indirectly, design feature preference is influenced by individual differences, such as user motivation, personality, and perceived Gamefulness (as demonstrated in the previous studies). Current measurements of user engagement vary in advantages, with self-report measures offering greater convenience but less objectivity, whereas task-performance measures offering greater objectivity but less convenience. An optimised measure that could build on the strengths of self-report and task-performance would be to make use of online free to play video games, which provide an environment in wherein objective engagement can be recorded without the resource intensiveness of conventional task-performance measures. This targeted form of user engagement can provide a more nuanced focus to how the effect of user design feature preference might correspond to objective forms of user engagement that will offer a stronger and more reliable insight for how Gamification can be more effectively adapted.

7.4 Research aim

The primary aim for the present study was to utilise a more objective measurement of user engagement that could further validate the significant relationships found in Study two between preference for EIA design features and variance in user engagement. As the measurement of engagement was to observe user interaction with online free to play video games, the design feature dimensions that could be assessed in relation to user engagement were limited (as not all design feature dimensions could be adequately mapped to an online free to play video game). Thus, of the EIA dimensions, only Expression and Improvement design features were assessed. It was expected that higher levels of preference for Expression and Improvement design features would correspond to higher levels of user engagement, similar to that which was observed in the previous study.

7.5 Methods

7.5.1 Design

The present study employed a within-subjects cross-sectional online methodology to identify how a task-performance measure of user engagement would relate to design feature preference. Participants were instructed to play a sequence of five online browser games (each different based on the design features they are comprised of) ad libitum and provide preference ratings for a series of design feature vignettes. As such there were two variables of interest; the time spent playing each of the five online browser games (measured in seconds) which represented user engagement; and responses to the DFPS, which measured user preference for 47 design features.

7.5.2 Participants

Players were recruited from various gaming Reddits; /r/Borderlands, /r/GamePhysics, /r/Overwatch, /r/Rainbow6, /r/Skyrim, /r/Steam, and /r/TrueGaming. A total of 325 users initially responded to the survey of which 72 users participated to completion. 90% of participants were male, 7% female, and 3% not identifying. Ages ranged from 18 to 77 (Mean age = 26.22 years; SD = 9.75). The average level of gaming experience for the sample was 18 years. 25% of participants were employed, 42% were students, 17% were actively seeking employment, and 4% were unemployed. Of the total sample, 40% were from North America, 40% from Europe, 9% from Asia, and the remaining from South America, Oceania, or Africa. 86% of participants reported most use of a PC platform to play video games, while the remaining 14% of participants reported most use of a console platform to play video games (e.g., XBOX, Playstation, Nintendo).

7.5.2.1 Eligibility criteria

Participants were only eligible to participate if they met two main criteria. The first, was that participants must have regularly played video games for at least two hours per week, which in previous gaming research has been regarded as the minimum time spent playing to qualify as being a gamer (Kolo & Braun, 2004). It was expected that players possess the prerequisite experience to understand the functional representation of Design features, therefore ratings of representativeness would be more reliable than those given by non-players. The second criteria were for participants to have been aged 18 years or

older. As the study exclusively employed online methodologies, there was no way for researchers to obtain parental consent for underage participants, therefore only those of the legal age to consent were able to participate.

7.5.2.2 Participation incentive

All participants were automatically enrolled into a free game giveaway, wherein the successful winner would be awarded a game of their choice, limited to $77 or £55 and purchasable only from a reputable online seller (such as STEAM, Origin, or the XBOX marketplace).

7.5.3 Materials

7.5.3.1 Design feature preference

Participants were required to complete the DFPS, which required participants to report to what extent they found a given design feature fun, motivating, useful, and preferable (Lopez & Tucker, 2019). A total of 47 design features were functionally represented via vignettes, each of which had been generated from a literature review (Arnab et al., 2015; da Rocha Seixa et al., 2016; Hamari & Lehdonvirta, 2010; Lameras, 2017; Marczewski, 2015; Nacke, 2018; Orji et al., 2018; Rocha et al., 2008; Sailer et al., 2017; Tondello et al., 2017; Werbach & Hunter, 2012), and previous study validation (see Table 16). Each vignette was limited to 17 words and achieved an average Flesch reading ease score of 53.31 and a Flesch-Kincaid grade level test score of 8.85, suggesting that the vignettes could be easily read by 13-14-year-olds. Together, all design features achieved an alpha of (a = .934) suggesting high internal consistency. Ratings were made via a 5-point Likert scale (1 – strongly unfavourable, 5 – strongly favourable) that was designed to emulate a conventional star rating system to induce participants to consider their ratings in a more meaningful capacity. Example vignettes include examples include the Complementarity design feature (being required to complete an objective or task only with the help and abilities of another player), and the Demotion design feature (being demoted and having your rank reduced after failing in some way).

7.5.3.2 Online games task

Participants were required to complete an online games task, during which they would play five online games consecutively, ad libitum. The list of games played can be found in Table 25 along with game descriptions. Each game was selected as per its design feature profile which was taken from the design feature factor loadings generated in Study 2. The range of games that could be selected was constrained due to needing to be free to play online, and not requiring substantial computing power to play. For example, the popular maze action game Pac-Man was selected as a Loss game, given its design feature profile corresponds to those of a Loss design feature dimension. To illustrate, the player controls Pac-Man, who must eat all dots inside the maze while avoiding the attacks of four ghosts' enemies. Within the maze, there are also power pellets, which once eaten enables Pac-Man to temporarily attack the ghosts. The power pellet feature is the equivalent to Item Degradation, which is a design feature characterised by a timed expiry of an item. Similarly, the Depletion design feature also exists within Pac-Man, as the user will lose the number of spare lives after each consecutive fail (three overall) until their progress resets to 0.

All games were accessible via the internet. Four out of five online game were hosted on the games website "Crazy Games", which serves as a licenced online games platform company that comprises of over 500 game developers, and works with several game studies (such as Ubisoft, Moonee, and Kiloo). Importantly, there were no requirements for significant computing power, meaning most users would not encounter any technical issues when playing the games. Additionally, there were no installation requirements of the participant to play the game. The online games task segment of the survey comprised of five sections. In each section, participants were presented with a URL that would direct them to the online game. The participants were instructed to play the online game until they were no longer interested, at which point they would return to the survey and proceed to the next section. The online games task was completed once all five sections (i.e., five games) had been played.

Table 25
Games selected for online games task

Design feature dimension	Game title	Game description	Design features present	
Expression	Online character creator	An online game wherein the user can customise and create a virtual character, with a significant degree of design choice. For example, being able to customise facial features, skin colour, and apparel.	Avatar, Design	2 out of 5
Improvement	Burnin rubber	A combat racing game in which you race against opponents and battle using weapons to achieve first place.	Item Power-up, Progress bar, Currency, Points	4 out of 7
Difficulty	Sonic Extreme Run	An obstacle game in which you control Sonic the Hedgehog to clearing each level, with the aim of accumulating as many gold rings as possible. You lose rings after each failed obstacle.	Behavioural momentum, Levels, Game Goal, Game Objective, Boss Battle	5 out of 5
Competition	Bullet force	A multiplayer first-person shooter wherein you face opponents within a game map, with the aim of defeating the opposing team using your weapons.	Leaderboard (feedback), Leaderboard (competition), PVP, Points, Rank	5 out of 7
Loss	Google Pacman	A maze action game wherein you control the Pac-Man. The aim is to eat all the dots placed within the maze while avoiding four colour ghosts. The level is complete once all dots are eaten.	Item Degradation, Depletion, Restriction	3 out of 6

7.5.3.3 User engagement

The metric of user engagement focused on how long participants spent playing each of the online games. As the online games were hosted on webpages external to the survey, a timer feature was used within the survey. Each game was presented to the participant on a separate page. When presented with the game, a timer feature would record the participant's first click on the page (following the URL for the online game) and the last click (where the page would be submitted, and the participant would then go to the next page). To encourage participants to access the online game, the option to proceed to the

next section did not become visible until after a 20 second buffer had passed. A score of engagement was calculated by the formula outlined in Figure 23.

Figure 23 - Formula for engagement buffers

Formula						
Submit page	-	First click	-	Buffer	=	Engagement score
Example						
50 seconds	-	10 seconds	-	20 seconds	=	20 seconds

7.5.3.4 Demographics

Participants were asked to submit standard demographic data, such as age, ethnicity, gender, education status, country of domicile, employment status and marital status. In addition, general gaming related demographics will also be collected, such as favourite game title, favourite game genre to play, most used gaming platform, and for how long they have played video games (in years).

7.5.4 Procedure

Participants accessed the survey via a URL link provided in the recruitment advertisement, after which they were presented with the study information sheet. After providing consent, participants completed the demographics sheet, followed by the design feature rating task, and the online game task (instructions for which can be found in Figure 24, Figure 25, and Figure 26), after which participants were finally debriefed.

Figure 24 - Initial briefing provided to participants for the online game task

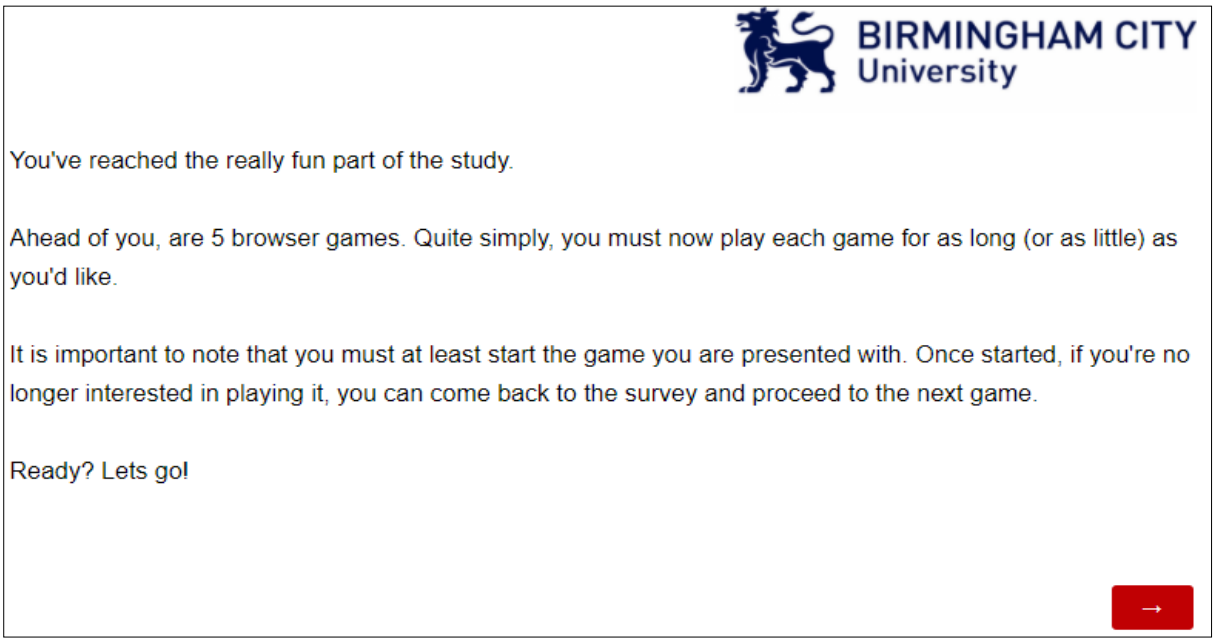


Figure 25 - The first page (of five) for online game task

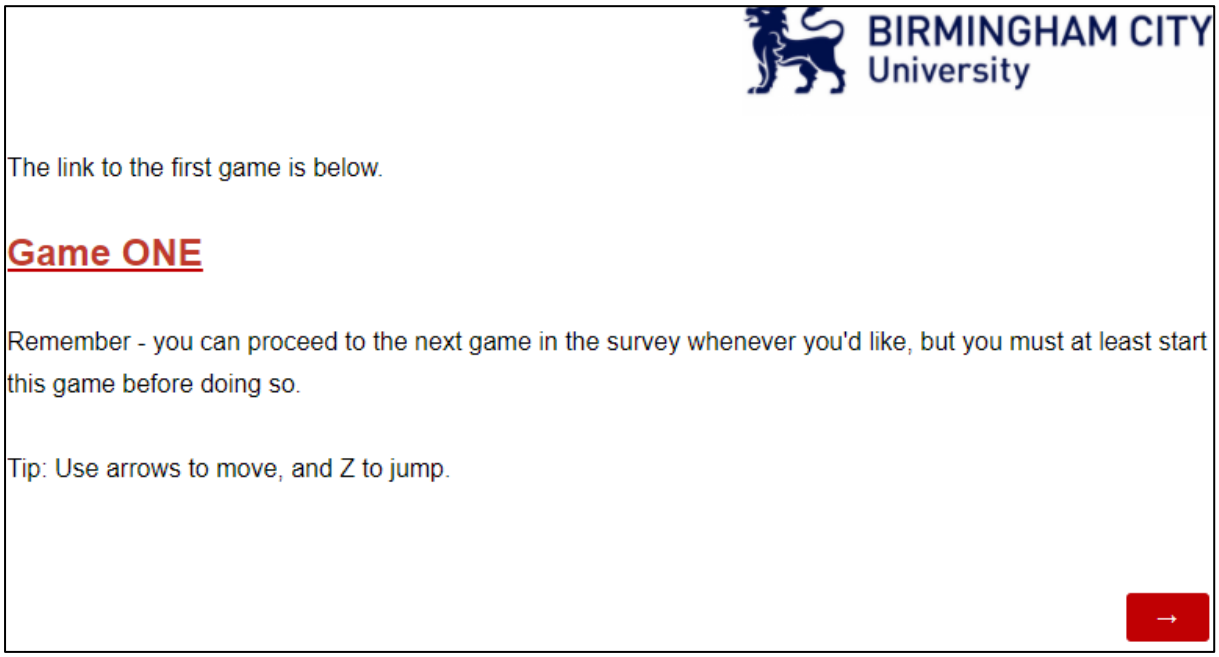
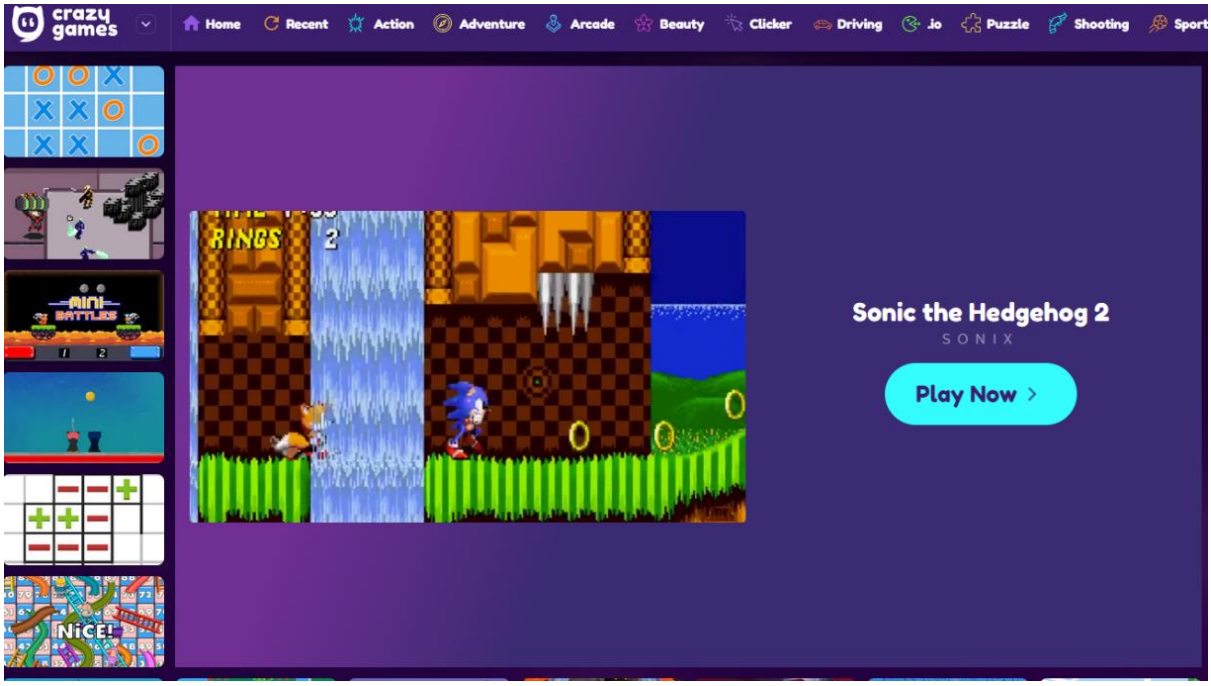


Figure 26 - The page to which participants would be redirected to following the Game URL



7.5.5 Ethics approval

Ethical approval was granted by Birmingham City University's research ethics committee under the reference code Lally/4978/R(A)/2020/Jan /BLSS FAEC

7.6 Results

A series of five multiple regressions were conducted to assess whether variance in user design preference could predict variance in user engagement when completing the online game task. Descriptive statistics showing the average duration for which each online game was played and design feature preference, as well as well as indications of normality (Skewness and Kurtosis), and the number of participants that provided responses can be found in Table 26.

Table 26

Distribution statistics and mean scores for user engagement and design feature preference

Variable	Mean	SD	Skew (SE)	Kurt (SE)	*n*
User engagement (time spent playing online browser games)					
Expression (Online Character Creator)	141s / 2.21m	244.84	3.84 (.283)	20.07 (.559)	72
Improvement (Burnin Rubber)	326s / 5.26m	1002.78	7.16 (.283)	55.96 (.559)	72
Difficulty (Sonic Extreme Run)	280s / 4.40m	611.17	4.45 (.283)	22.52 (.559)	72
Competition (Bullet Force)	194s / 3.14m	297.85	3.17 (.283)	13.08 (.559)	72
Loss (Google Pacman)	91s / 1.30m	137.02	3.87 (.283)	20.01 (.559)	72
Design features					
PVP	12.58	4.27	-0.04 (.283)	-0.58 (.559)	72
Leaderboard (Competition)	13.33	4.97	-0.60 (.283)	-0.61 (.559)	72
Punish	8.89	4.95	0.37 (.283)	-1.21 (.559)	72
Complementarity	11.88	4.61	-0.25 (.283)	-0.78 (.559)	72
Shared Goal	15.13	3.73	-0.58 (.283)	-0.12 (.559)	72
Trade	14.35	3.76	-0.47 (.283)	-0.17 (.559)	72
Friend Invite	17.67	4.05	-2.50 (.283)	6.68 (.559)	72
Text Chat	14.25	4.07	-0.50 (.283)	-0.27 (.559)	72
Voice Chat	13.68	5.08	-0.51 (.283)	-0.73 (.559)	72
Emotes	12.46	4.85	-0.19 (.283)	-0.93 (.559)	72
Communal discovery	15.71	3.74	-0.90 (.283)	0.72 (.559)	72
Trophy	15.19	4.13	-0.66 (.283)	-0.09 (.559)	72
Badges	14.14	4.51	-0.42 (.283)	-0.61 (.559)	72
Medal	14.00	4.60	-0.38 (.283)	-0.74 (.559)	72
Tokens	15.22	4.68	-0.97 (.283)	0.06 (.559)	72
Item Power-up	16.18	3.83	-1.42 (.283)	2.24 (.559)	72
Bonus	16.31	2.92	-0.71 (.283)	0.36 (.559)	72
Lottery	9.61	4.31	0.53 (.283)	-0.33 (.559)	72
Depletion	10.64	4.78	0.22 (.283)	-0.76 (.559)	72
Restriction	8.43	4.67	0.78 (.283)	-0.42 (.559)	72
Demotion	9.57	4.83	0.35 (.283)	-0.99 (.559)	72
Points	16.04	3.02	-0.41 (.283)	-0.70 (.559)	72
Progress Bar	16.49	2.97	-0.64 (.283)	-0.05 (.559)	72

Leaderboard (Feedback)	14.39	4.77	-0.64 (.283)	-0.43 (.559)	72
Scarlett letter	10.00	4.50	0.32 (.283)	-0.73 (.559)	72
Performance graphs	14.28	4.16	-0.53 (.283)	-0.30 (.559)	72
Walkthrough	12.44	3.91	0.09 (.283)	-0.64 (.559)	72
Tips / Hints	13.61	3.67	0.14 (.283)	-0.75 (.559)	72
Notification / Prompts	10.58	4.37	0.10 (.283)	-0.93 (.559)	72
Cut Scenes	14.96	3.77	-0.55 (.283)	-0.07 (.559)	72
Storyline	17.29	3.33	-0.95 (.283)	-0.19 (.559)	72
Currency	15.90	4.02	-0.75 (.283)	-0.23 (.559)	72
Item Degradation	9.89	5.13	0.52 (.283)	-0.95 (.559)	72
Dashboard	15.67	3.61	-0.69 (.283)	0.35 (.559)	72
Behavioural Momentum	16.29	3.78	-1.23 (.283)	1.85 (.559)	72
Levels	14.57	4.24	-0.82 (.283)	0.18 (.559)	72
Barriers / Access	10.57	4.61	0.11 (.283)	-0.98 (.559)	72
Game Goal	15.90	3.35	-0.92 (.283)	1.20 (.559)	72
Game Objective	16.69	3.23	-1.17 (.283)	2.06 (.559)	72
Boss Battles	16.83	3.15	-1.13 (.283)	2.27 (.559)	72
Beginners luck	12.17	4.36	-0.18 (.283)	-0.44 (.559)	72
Design / Editing / Customisation	16.26	4.02	-1.10 (.283)	0.81 (.559)	72
Decision Making	17.47	3.09	-1.12 (.283)	0.86 (.559)	72
Vote	14.53	4.48	-0.49 (.283)	-0.51 (.559)	72
Avatar	14.92	4.58	-0.36 (.283)	-1.10 (.559)	72
Profile	14.14	5.02	-0.59 (.283)	-0.59 (.559)	72
Rank / Status	14.43	4.52	-0.31 (.283)	-0.73 (.559)	72

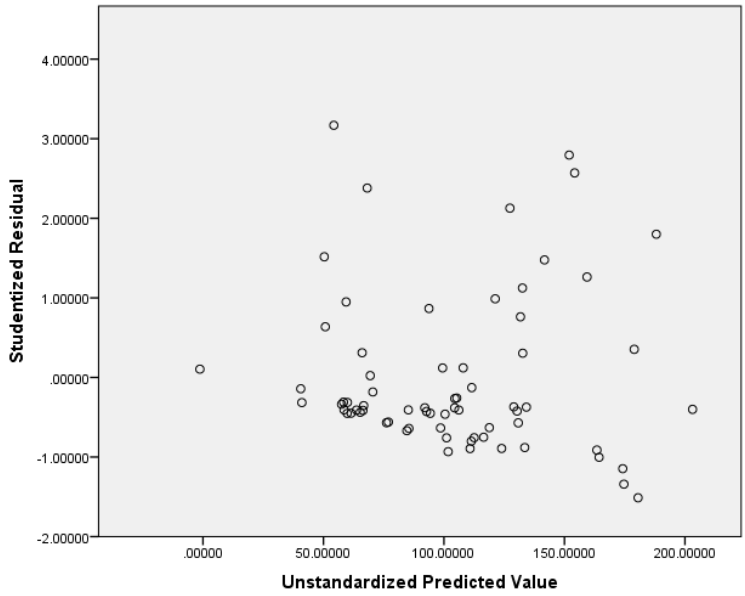
Notes: Mean score for user engagement is presented in the format of 'seconds / minutes', SD = Standard deviation, Skew = Skewness, Kurt = Kurtosis, SE = Standard error, n = Sample size.

7.6.1 Design feature preference and engagement with an Expression game

A multiple regression was conducted to assess to what extent user design preference predicted user engagement with an online Expression game (Online Character creator). There was linearity as assessed by partial regression plots and a plot of studentized residuals against the predicted values. There was independence of residuals as assessed by a Durbin-Watson statistic of 2.178. There was homoscedasticity as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values (Figure 27). There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1 (Hair et al., 2014), indicating that the independent variables (design feature dimensions) were not highly correlated with each other. Measures to detect outliers revealed no studentized deleted residuals greater than ±3 standard deviations (Everitt & Skrondal, 2010), no leverage values greater than 0.2 (Huber, 1981), or values (influential points) for Cook's distance above 1 (Cook, 1977), indicating no outliers were

present in the data. The assumption of normality was met, as assessed by visual inspection of a Q-Q Plot (Appendix Q). The multiple regression model was not significant in predicting user engagement with an Expression game, $F(9, 59) = .653$, $p > .050$, adj. $R^2 = .05$.

Figure 27 – Plot of design feature preference and engagement with an Expression game studentized residuals

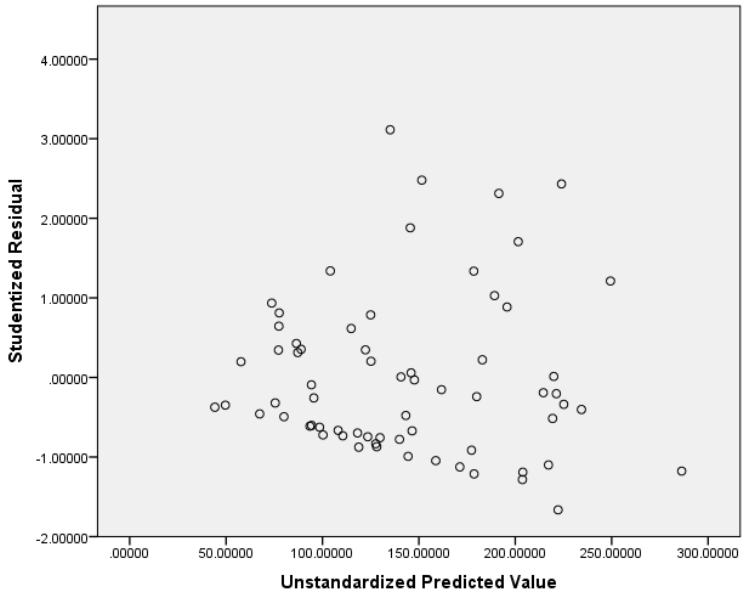


7.6.2 Design feature preference and engagement with an Improvement game

A multiple regression was conducted to assess to what extent user design preference predicted user engagement with an online Improvement game (Burnin Rubber). There was linearity as assessed by partial regression plots and a plot of studentized residuals against the predicted values. There was independence of residuals as assessed by a Durbin-Watson statistic of 1.670. There was homoscedasticity as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values (Figure 28). There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1 (Hair et al., 2014), indicating that the independent variables (design feature dimensions) were not highly correlated with each other. Measures to detect outliers revealed no studentized deleted residuals greater than ±3 standard deviations (Everitt & Skrondal, 2010), no leverage values greater than 0.2 (Huber, 1981), or values (influential points) for Cook's distance above 1 (Cook, 1977), indicating no outliers were present in the data. The assumption of normality was met, as assessed by visual inspection of a Q-Q Plot (Appendix R). The multiple regression model was not significant

in predicting user engagement with an Improvement game, $F(9, 57) = .862$, $p > .050$, adj. $R^2 = .02$.

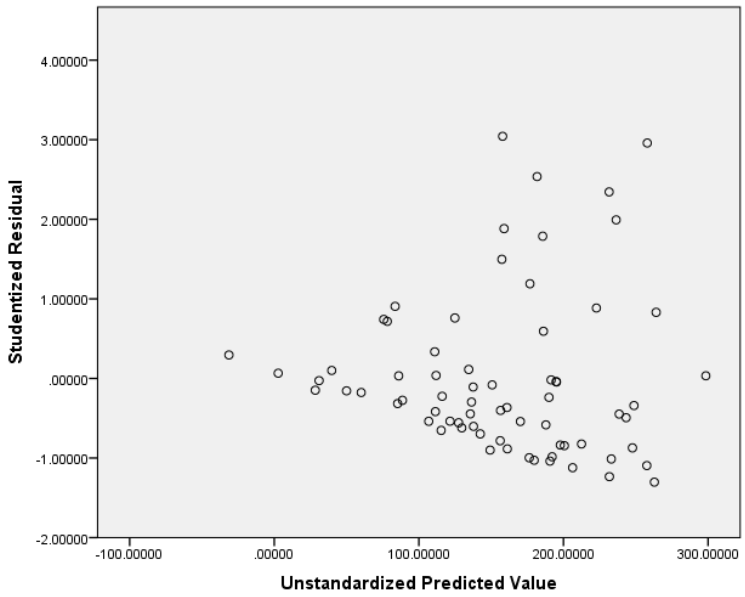
Figure 28 – Plot of design feature preference and engagement with an Improvement game studentized residuals



7.6.3 Design feature preference and engagement with a Competition game

A multiple regression was conducted to assess to what extent user design preference predicted user engagement with an online Competition game (Bullet Force). There was linearity as assessed by partial regression plots and a plot of studentized residuals against the predicted values. There was independence of residuals as assessed by a Durbin-Watson statistic of 1.670. There was homoscedasticity as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values (Figure 29). There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1 (Hair et al., 2014), indicating that the independent variables (design feature dimensions) were not highly correlated with each other. Measures to detect outliers revealed no studentized deleted residuals greater than ±3 standard deviations (Everitt & Skrondal, 2010), no leverage values greater than 0.2 (Huber, 1981), or values (influential points) for Cook's distance above 1 (Cook, 1977), indicating no outliers were present in the data. The assumption of normality was met, as assessed by visual inspection of a Q-Q Plot (Appendix S). The multiple regression model was not significant in predicting user engagement with an Competition game, $F(9, 60) = .931$, $p > .050$, adj. $R^2 = .01$.

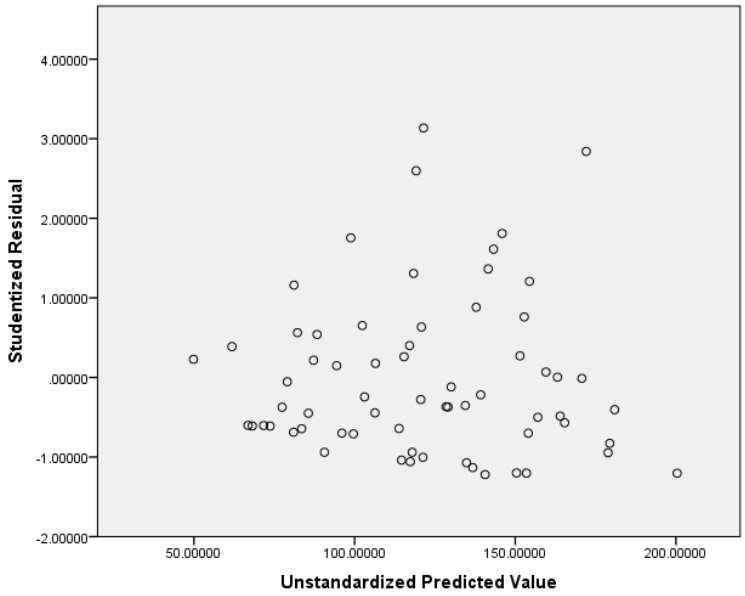
Figure 29 – Plot of design feature preference and engagement with a Competition game studentized residuals



7.6.4 Design feature preference and engagement with a Difficulty game

A multiple regression was conducted to assess to what extent user design preference predicted user engagement with an online Difficulty game (Sonic Extreme run). There was linearity as assessed by partial regression plots and a plot of studentized residuals against the predicted values. There was independence of residuals as assessed by a Durbin-Watson statistic of 1.670. There was homoscedasticity as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values (Figure 30). There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1 (Hair et al., 2014), indicating that the independent variables (design feature dimensions) were not highly correlated with each other. Measures to detect outliers revealed no studentized deleted residuals greater than ±3 standard deviations (Everitt & Skrondal, 2010), no leverage values greater than 0.2 (Huber, 1981), or values (influential points) for Cook's distance above 1 (Cook, 1977), indicating no outliers were present in the data. The assumption of normality was met, as assessed by visual inspection of a Q-Q Plot (Appendix T). The multiple regression model was not significant in predicting user engagement with an Competition game, $F(9, 60) = .931$, $p > .050$, adj. $R^2 = .01$.

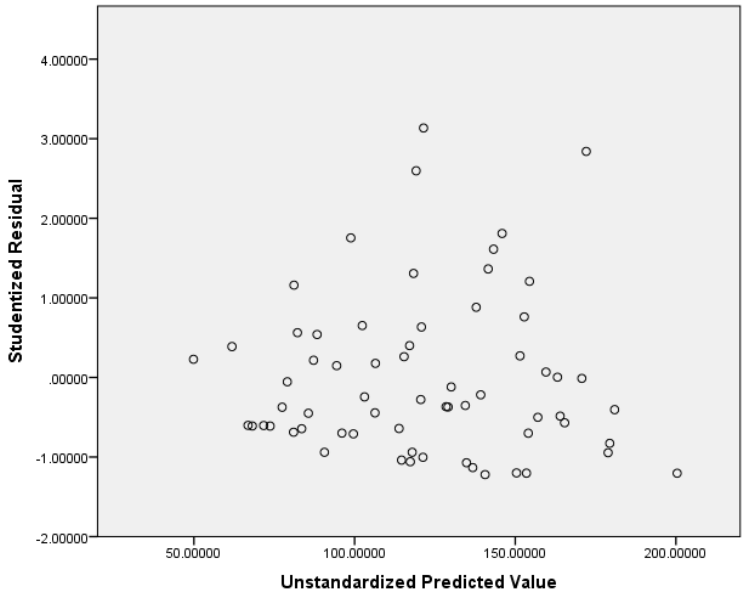
Figure 30 – Plot of design feature preference and engagement with a Difficulty game studentized residuals



7.6.5 Design feature preference and engagement with a Loss game

A multiple regression was conducted to assess to what extent user design preference predicted user engagement with an online Loss game (Google Pacman). There was linearity as assessed by partial regression plots and a plot of studentized residuals against the predicted values. There was independence of residuals as assessed by a Durbin-Watson statistic of 1.670. There was homoscedasticity as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values (Figure 31). There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1 (Hair et al., 2014), indicating that the independent variables (design feature dimensions) were not highly correlated with each other. Measures to detect outliers revealed no studentized deleted residuals greater than ±3 standard deviations (Everitt & Skrondal, 2010), no leverage values greater than 0.2 (Huber, 1981), or values (influential points) for Cook's distance above 1 (Cook, 1977), indicating no outliers were present in the data. The assumption of normality was met, as assessed by visual inspection of a Q-Q Plot (Appendix U). The multiple regression model was not significant in predicting user engagement with an Competition game, $F(9, 60) = .931$, $p > .050$, adj. $R^2 = .01$.

Figure 31 – Plot of design feature preference and engagement with a Loss game studentized residuals



7.7 Discussion

The primary aim of the present study was to utilise a more objective measurement of engagement (task-performance) that could further validate significant relationships between user engagement and design feature preference found in Study two. Specifically, whether variance preference for EIA design feature dimensions would correspond to significant variance in observed user engagement. It was hypothesised that of the significant relationships found in the previous study that were retested in the present study, would also be found as significant. Of the five relationships examined in the present study (Table 27), the relationship between user engagement and preference for Expression and Improvement design features were retests, whereas the relationship between user engagement and preference for Loss, Difficulty, and Competition design features were being examined for the first time. The results indicated that all relationships were found to be non-significant, therefore rejecting the hypothesis. Specifically, preference for Expression design features (Avatar; and Design), Improvement design features (Item Power-up; Progress Bar; Currency; and Points), Difficulty design features (Behavioural momentum; Levels; Game Goal' Game Objective; and Boss Battle), Competition design features (Leaderboard-Feedback; Leaderboard-Competition; PVP; Points; and Rank), and Loss design features (Item Degradation; Depletion; and Restriction) was not significantly associated with variances in user engagement.

Table 27
Relationships examined in the present study

Design feature dimension	Examined in Study two
Expression	Yes
Improvement	Yes
Difficulty	No
Competition	No
Loss	No

The difference in results between Study two and present study could be explained by a variety of factors. One example is the role of computer graphics and cognitive complexity

(Sedig et al., 2017). Research indicates that user engagement can often be determined by two design considerations; representation (how the game is visually displayed) and interaction (how the user is enabled to make use of the game information) (Haworth et al., 2010). The scope of these design considerations can be largely determined by the computing power afforded to game designers, with greater computing power providing a more sophisticated level of design (i.e., high end graphics and more game complexity). In the present study, participants were instructed to play online browser games, which when compared to PC or console games, are limited in the computing power used for game design and game operation. One could argue that such differences result in a varied user experience, with more sophisticated games being more engaging, and less sophisticated games being less engaging. On this basis, it could be expected that users who might be used to PC or console games would find online browser games less engaging, given the differences in computing power and subsequent differences in game graphics of complexity.

The differences in the user experience between online browser games and PC/console games might have had a marked influence on user engagement and explain why the outcomes of Study two were not replicated with the present study. First, all participants in the present study exclusively reported preferring to play video games on either a PC or a games console (see Table 28).

Table 28
Preferred video game platform

Platform	***n***
PC	59
XBOX	5
Playstation	4
Nintendo	1

All participants were also recruited from Reddits that focused on either PC games or console games, indicating a level of interest in these games beyond playing (given the role of the Reddits are to facilitate game discussion) (see Table 29).

Table 29
Reddits from which participants were recruited

Reddit	***n***
Skyrim	37
Gamephysics	12
Steam	12
True gaming	6
Overwatch	3
Borderlands	1

Finally, participants were asked to report their favourite video game of all time, to which most responded with The Elder Scrolls V: Skyrim (2011): a game that received critical acclaim for its graphical fidelity. Other games reported (Table 30) were also titles that can only be played on PC or a games console (and not as an online browser game). Therefore, it could be argued that due to the characteristics of the participant sample preferring to play PC or console games, the difference in user experience when playing an online browser game could have resulted in a reduced level of user engagement.

Table 30
Reddits from which participants were recruited

Game title*	***n***
Skyrim	11
Witcher	5
Minecraft	4
Mass effect	3
Bioshock Infinite	2
GTA V	2
Hollow Knight	2

Notes: *game titles which received at least two or more mentions are included above. In total there were 48 unique titles, all of which can be found in Appendix V.

Furthermore, the issue also exists that online browser games are limited in their design and depth, such that they do not comprise of a complete design feature profile. For example, the design feature dimensions tested in the present study collectively comprised of 30 design features, with the corresponding online browser games only featuring 19 design features of interest. Although an online browser game does not necessarily need to comprise of all design features to be engaging, in the context of the present study and further development of the DFPS, a complete profile of design features being featured in a given game is highly desirable. Given the greater capabilities of PC and console games, it could be expected that such games are more likely to comprise a complete design feature profile, and therefore a more appropriate research tool in the context of this project. Further development of the DFPS in relation to observations of user engagement might be better suited if PC or console games are included instead of online browser games.

At a closer look, measurements of user engagement might also be refined in future work. The present study employed a measurement of user engagement by using the time spent completing the online game task (i.e., participants playing a given game). Although this approach mitigated the subjectivity associated with self-report measures, the time spent completing the online game task was still not directly observable and instead indirectly calculated (time taken between first and last click). As such, recorded user engagement could not be directly verified. In comparison, if the online game task were to be completed in-person within an offline laboratory setting, user engagement could be directly observed, which addressed both the issue of indirectly calculating user engagement and assist with recorded engagement being verifiable. In the context of the present study, this approach was preferred and originally planned (with research ethics also being approved). However, due to the Covid-19 pandemic and subsequent limitations on social interaction, the university had issued a policy that prevented face to face data collection for an extensive length of time, meaning this approach could not be used.

Another explanation for why engagement was not associated with design preference as expected could be the frequency of game change during the online ask. Prior to beginning the task, participants were advised that there would be several games to play, which might have inadvertently primed participants to move to the next game sooner. Such behaviour could be explained by the cognitive psychology concept of overchoice, which

refers to the consequential impact on engagement and satisfaction when faced with numerous options (Misuraca et al., 2016) . Research indicates that when presented with significant variety, individuals become less satisfied with the options that are available, irrespective of what they choose (Gourville & Soman, 2005). In the context of the online game task, the disclaimer that there will be five games to play might have served as a primer for participants to become dissatisfied with their play experience, and consequently reduced their level of engagement. However, the role of overchoice often concerns instances where there exist options to choose between. The present study did not enable participants to choose between which game they played, therefore this concept might not fully explain the inconsistency in results.

In the context of understanding how Gamification could be improved, methods which expose the user to multiple games may not be suitable in emulating a natural Gamification environment, given that in a Gamification context a user is unlikely to interact with several systems in a consecutive format, or interact with several systems that differ in aesthetics, function, or purpose. A different approach to measuring engagement that would mitigate the issue of indirect engagement observation and the repeat exposure to multiple games, would be to focus on in-game behavioural metrics. For instance, the present study calculated user engagement by indirectly observing how long a player spent playing an online browser game. By comparison, in-game behavioural metrics, such as number of matches played, number of online friends, and number of hours spent playing a given game mode, could provide a more direct and objective indication of user engagement. In-game behavioural metrics are arguably more representative of user engagement, as they indicate natural player behaviours and tendencies outside of the experimental context, from which a more representative picture of overall user engagement could be painted. In comparison, user engagement that is observed within an artificial context might certainly be representative of engagement in some capacity, but that what is observed might not correspond to contexts outside of the artificial setting. Research indicates that due to the artificial nature of experimental designs, participant behaviour may not consist with behaviour performed in more realistic settings, given that experimental research can create situations that are not realistic (Kachelmeier & Towry, 2015). As such, metrics of engagement that build upon natural play behaviour might be more reliable and possess a far-reaching continuity between research and research application.

7.8 Summary

The primary outcome of the present study was to assess whether previously identified relationships between design feature preference and user engagement would be replicated when using a task-performance measurement of user engagement. The results revealed that none of these relationships were consistent, such that reported user design feature preference did not predict or coincide with the time each user spent playing an online browser game. Explaining these outcomes, the present study highlighted several methodological limitations of the study design used. These design limitations might be improved if alternative measurements of user engagement can be used within further research. One such example would be to measure user engagement via in-game player activity, making use of objective in-game player metrics. Despite these limitations, the present study still made significant contributions to the area of Adaptive Gamification, such as highlighting online browser games as being an ineffective method at curating natural player behaviour. Furthermore, although non-significant relationships were found between design feature preference and reported engagement, these findings do not necessarily indicate a limitation in the DFPS, given that its primary function is to measure user design feature preference, which in the context of this study, was achieved.

8. Study Five

8.1 Abstract

Although a task-performance measurement of user engagement improves on the limitations associated with self-report measurements, observed player behaviour is arguably only representative of the given context within which it is observed. As such, the degree to which a "snapshot" of player behaviour can predict the users' natural play behaviour is arguable. In-game behavioural metrics, which aggregate and track player behaviour over time and source player data from genuine game activity have been used in previous research and are in several ways more representative of natural play behaviour. As part of further validation of the DFPS, the present study sought to test previously identified relationships between user design feature preference and user engagement as measured by in-game behavioural metrics. Using a cross-section online design, 24 League of Legends players completed the DFPS and provided a range of activity and in-game behavioural statistics (e.g., rank, player level, points, items owned). Correlation analysis revealed that none of the previously identified relationships between user design feature preference and user engagement were found to be significant. Methodological explanations and further research is discussed, such as the value of using server-side data to provide in-game behavioural data, as well as the potential of using purpose-built in-game behavioural metrics.

8.2 Introduction

As levels of user engagement indicate whether Gamification has been successful, effective measurements of engagement are of fundamental importance to the development and improvement of Gamification. In the previous study, the strengths and weaknesses of task-performance measurements of engagement were contrasted with conventional self-report measures. It was determined that a measurement which optimises the objectivity of real-time task-performance but incorporates the ease of self-report for participants could be an effective alternative. The findings were not consistent with those from earlier studies in the project. In review of the methodologies used, a case was made that such inconsistencies could be due to indirectly calculating user engagement from participants playing online browser games, in addition to measuring user engagement in an artificial setting. In response, it was suggested that a more effective measurement of user engagement could be found in analysing in-game behavioural metrics (e.g., number of matches played, number of online friends, and number of hours spent playing a given game mode), which provide a potentially more representative indication of user engagement, given they reflect natural player behaviours and tendencies over time.

Existing gaming research has adopted a similar approach to understanding user engagement by looking toward in-game user behaviour (i.e., how the user plays the game and their overall play activity). For instance, in examining user motivation, Billieux et al. (2013) compared self-reported motives with in-game behaviour to substantiate whether there was continuity between self-report player tendencies and their natural play behaviour. The study found that specific associations such as motives for teamwork and competition being strong predictors of in-game advancement within World of Warcraft (a team-based Massively Multiplayer Online Role-Playing Game). Similarly, Kahn et al. (2015) reported additional in-game data to further substantiate the motivational tendencies of participants. For example, participants who reported higher socialiser motives also demonstrated more in-game social behaviours, such as playing with larger teams, and possessed a greater degree of in-game social capitol, such as more online friends with whom they played with more regularly. Additionally, users who reported greater game completion tendencies also had played with a higher number of champions (characters with different abilities that the user could select and use in a game mode). Given the characteristics of a completionist motivation, which is to work towards

completing all objectives and tasks that exist within the game, the in-game user behaviour consisted with the self-reported completion motives. In addition, Yee et al. (2011) found correlations between personality traits and in-game behaviour, with higher levels of extraversion being associated with a more frequent completion of challenges that required teamwork, while lower extraversion was associated with more frequent completion of solo challenges.

Compared with self-report (Study 2-3) and experimental task-performance (Study 4) measurements of user engagement, a shift to measuring engagement by in-game behavioural metrics could provide more reliable results on the basis of three key propositions: in-game behavioural metrics provide a higher degree of ecological validity and predictive value, given they capture primary player behaviour as it occurs over time (Yee et al., 2012); in-game behavioural metrics are unlikely to be skewed by the initial time taken by users' to onboard or familiarise with a game; and in-game behavioural metrics are more objective.

8.2.1 Higher ecological validity

A limitation of experimental research (Study 4) is the artificial environment that is created for the participant (Lahti, 2015), such that behaviour observed in the artificial environment often will not consist with behaviour that occurs in natural environments, given the changes in environmental conditions. In the context of Gamification research, user engagement derived from experimental research (Study 4) could be characterised by artificial or short-term engagement patterns, which may not persist beyond the experimental environment.

Furthermore, experimental research similar to the previous study (wherein a variety of design features are presented to a participant) might also provide unreliable indications of user engagement, mainly due to the novelty effect. In asking participants to interact with new design features not previously interacted with, the response in user activity could be characterised by spikes of engagement due to the novelty of such interactions, as opposed to underlying design feature preference. Research has found that high levels of activity are often observed immediately when a user interacts with new design features, however, their activity levels drop shortly after and the novelty has worn off

(Kofinas & Tsay, 2019). As such, in-game behavioural metrics might provide a more representative account of user engagement.

8.2.2 Distinguishing between processes of familiarisation and onboarding

Another issue with experimental measurements of user engagement is the difficulty in separating between time spent familiarising oneself with the game environment and controls, and time spent playing due to enjoyment and genuine interest in the game (i.e., engagement). To illustrate, many games introduce and inform a new player about how they can operate in the game world by way of textual instructions (see Figure 32).

Figure 32 – textual instructions (Death Stranding, 2019)

One can argue that the time spent understanding the game in this way is not indicative of the type of user engagement that can be leveraged and replicated through the design of Gamification. Instead, this process of familiarisation could be considered as one of several different preliminary steps that prepare the user to engage with the game as intended by the game designers. Moreover, the time spent becoming accustomed to a game depends on the skill and aptitude of the user, which can often vary (Basak et al., 2011), therefore distinguishing between time spent familiarising and playing is made more difficult.

In comparison, measuring engagement via in-game behavioural metrics would mitigate the need to distinguish whether any engagement recorded comprised of the user becoming familiar with the game, or a product of their enjoyment. For example, the

measurement used in the previous study comprised of a cross-sectional assessment of user engagement, such that level of engagement was determined on the basis of a single play session. Considering the previously discussed novelty effect, as well as the role of familiarisation, engagement calculated from a single play session is liable to environmental influences, such that player activity and patterns may only be exhibited in the experimental environment.

By comparison, play which occurs over time in a natural play environment is arguably more representative of genuine user engagement, given that the player has had enough time to become accustomed to the game, and focus or orient their play towards what they enjoy (and not what they encounter during a single play session). As different behavioural metrics would be used to provide an overall account of user engagement (such as number of games played, or number of coins accumulated), a more nuanced indication of user engagement is also provided, as opposed to the umbrella metric of time spent playing. Of notable mention, is that in-game behavioural metrics would also facilitate clear distinguishment between time spent familiarising with the game and time spent enjoying the game. For instance, metrics which possess intrinsic ranking systems, wherein more desirable versions of something are rewarded or acquired only after in-game milestones are achieved. One such example are tiered trophies, with bronze trophies being awarded first, followed by silver trophies and finally gold trophies, each of which become more difficult to acquire over time. A user who owns the silver or gold trophies would have likely dedicated more time playing the game to acquire the more valuable trophies, given that acquisition of these items requires more time spent playing, or more time spent developing skills needed to complete the necessary challenges to acquire them (Gurwin, 2021). However, it must be noted that intrinsic ranking systems may not necessarily be best observed in the acquisition of in-game items, as the means to acquire the item (such as in-game currencies accrued over time) are not exclusively accumulated by time spent playing alone. For instance, some users may acquire in-game items via in-game purchases or gifting from other players.

8.2.3 Highly objective – recollection vs actual play tendencies

In comparison to self-report measures of engagement that require the participant to recall their level of engagement (e.g., how long do you play for?) (Brunborg et al., 2014), in-game behavioural metrics provide a greater degree of objectivity. According to Lucas

(2018), self-report measures rely on a user's subjective evaluation of their behaviour, which may not provide an accurate representation of what is being reported (Demetriou et al., 2015). Moreover, in the context of gaming, research has found that players can often underreport play activity (Jeong et al., 2018). In determining user engagement via in-game behavioural metrics, researchers would be provided with player statistics that are accurately recorded and automatically updated in real-time, thereby rendering participant recall unnecessary.

For instance, the popular MOBA developed and published by Riot Games', *League of Legends* (2009) provides users with profile dashboard, within which users are provided with data from several in-game behavioural metrics, such as overall wins, losses, earned and spent points, and so forth. Quantified data of this type is aggregated and updated in real-time and is calculated by in-game monitoring of user activity which is then associated with the user's game account. In addition to providing more accurate and objective indications of player activity, in-game behavioural metrics are also likely to provide a more representative profile of user engagement. To illustrate, whereas self-report measures would require participants to estimate their average play duration (Study 2 and 3), in-game behavioural metrics, such as overall number of games played (League of Legends, 2009) would provide a more accurate indication of time spent playing, given that one can calculate play duration based on how many games have been completed (by way of referencing average match durations detailed in Table 31; taken from "League of Graphs", a League of Legends statistics engine).

Table 31
Average duration of LoL games by user rank (Çakır, 2021)

User rank	**Average game duration in minutes**
Iron	29:44
Bronze	29:40
Silver	29:32
Gold	29:08
Platinum	28:24
Diamond	27:13
Master	26:20

GrandMaster	25:15
Challenger	24:33

Furthermore, given the quantified nature of in-game behavioural metrics, further analyses and relationships can also be more reliably identified and explored. For example, research that has focused on the role of video game play and prosocial behaviour has often employed self-report measures (Gentile et al., 2009; Ihori et al., 2007), with participants asked to reflect on how often they were kind to other people (Ihori et al., 2007), or how often they would dedicate time to help others (Gentile et al., 2009). A persisting limitation with self-report measures is that participants may not correctly recall and report previous behaviour (Demetriou et al., 2015), and in the context of prosocial behaviour, studies have indicated that when asked about prosociality response bias can take the form of impression management, wherein participants report in a way they perceive will make them appear more favourably to the researcher (Sassenrath, 2019).

In comparison, in-game behavioural metrics may provide a more robust evidence base (Johannes et al., 2021). For example, prosocial gaming behaviours such as cooperation and sharing (Mengel, 2014) could be measured via in-game behavioural metrics, such as how often a user plays in a team in comparison to playing solo (Kahn et al., 2015) or how often a user gifts items to others. In using this data, a more reliable metric of prosociality can be calculated and thus a more objective assessment of how video game play influences participant prosocial behaviour could be enabled.

8.2.4 Substantiate user design preference

In-game behavioural metrics also serve to substantiate the developed measure of user design feature preference. As users are asked to rate to what extent they prefer design features, such preference ratings can be validated by triangulating with the users in-game behaviour. Doing so would provide indication on whether there exists continuity between what the user reportedly prefers and their in-game activity.

8.3 Rationale

Compared with self-report measures that rely on recall and experimental task performance measures that produce artificial engagement, in-game behavioural metrics provide a higher degree of objectivity and representativeness of the user's natural play activity, which is automatically aggregated in real-time. Importantly, these metrics can further validate the DFPS, given that reported preference can be triangulated with in-game activity. As such, the present study aims to further test relationships between user engagement and user design feature preference by using in-game behavioural metrics as an alternative measurement of user engagement.

8.4 Research aim

The primary aim for the present study was to use in-game behavioural metrics as a measurement of user engagement and assess the predictive value of the DFPS, by assessing whether user design preference would correlate with reported in-game behaviour. It was expected that user design preference would correlate with a corresponding in-game behavioural metric of user engagement.

8.5 Methods

8.5.1 Design

The present study employed a within-subjects cross-sectional online methodology to identify how an in-game behavioural metric measure of user engagement would relate to design feature preference. Participants were instructed to complete an in-game behavioural metric scale and provide preference ratings for a series of design feature vignettes. As such there were two variables of interest; responses to the in game behavioural metric scale, which was adapted the MOBA game *League of Legends* (2009) that would record player statistics (accessed via the League of Legends user account dashboard); and responses to the DFPS, which measured user preference for 47 design features.

8.5.2 Participants

Players were recruited from League of Legends Reddits (/r/LeagueOfLegends, /r/LeagueOfLinux, and /r/LeagueConnect) and discord servers (/r/LoLeSports, /r/LeagueOfLegends, and /r/LeagueConnect). A total of 180 users initially responded to the survey of which 24 users participated to completion. 87% of participants were male while 13% were female. Ages ranged from 18 to 29 (Mean age = 21.20 years; SD = 3.00). The average level of gaming experience for the sample was 9.04 years. 23% of participants were employed, 4% were students, 43% were actively seeking employment, and 4% were unemployed. Of the total sample, 35% were from North America, 52% from Europe, 4% from Asia, and the remaining from South America, Oceania, or Africa.

8.5.2.1 Eligibility criteria

The eligibility criteria for Study five comprised of three elements, two of which were the same as the previous four studies (playing for a minimum of two hours per week, and being aged 18 years or older), and the third being for participants to spend most of their gaming time playing League of Legends. It was expected that participants who met this third condition would possess the necessary understanding to continue with the in-game behaviour scale. For instance, understanding terms such as “draft picks”, “riot points”, and “blue essence”.

8.5.3 Materials

8.5.3.1 Design feature preference

Participants were required to complete the DFPS, which required participants to report to what extent they found a given design feature fun, motivating, useful, and preferable (Lopez and Tucker, 2019). A total of 47 design features were functionally represented via vignettes, each of which had been generated from a literature review (Arnab et al., 2015; da Rocha Seixa et al., 2016; Hamari & Lehdonvirta, 2010; Lameras, 2017; Marczewski, 2015; Nacke, 2018; Orji et al., 2018; Rocha et al., 2008; Sailer et al., 2017; Tondello et al., 2017; Werbach & Hunter, 2012), and previous study validation (see Table 16). Each vignette was limited to 17 words and achieved an average Flesch reading ease score of 53.31 and a Flesch-Kincaid grade level test score of 8.85, suggesting that the vignettes could be easily read by 13-14-year olds. Together, all design features achieved an alpha

of (a = .934) suggesting high internal consistency. Ratings were made via a 5-point Likert scale (1 – strongly unfavourable, 5 – strongly favourable) that was designed to emulate a conventional star rating system to induce participants to consider their ratings in a more meaningful capacity. Example vignettes include examples include the Complementarity design feature (being required to complete an objective or task only with the help and abilities of another player), and the Demotion design feature (being demoted and having your rank reduced after failing in some way).

8.5.3.2 In-game behavioural scale

Participants were required to complete a 14-item in-game behaviour scale which was adapted to *League of Legends* (2009) and assessed key player behaviour. Of the 14-items, nine required participants to consult the League of Legends player dashboard, which were referred to as analytic items. The remaining six items required participants to consider their recent player behaviour and were referred to as reflective items. Example items from the analytic dimension included "What is your summoner level?" and "What is your total mastery score?". Example items from the reflective dimension included "How often do you check the leaderboards every week?" and "How many chests and capsules did you receive last month?". All metrics measured by the in-game behavioural scale were also mapped onto the user design features for which participants were to provide preference ratings for. For example, the analytic item "What is your summoner level?" would map onto the Rank design feature, which is defined as "being assigned a category and rank that reflects your ability, score, and/or experience". A full summary of the in-game behaviour scale can be found in Table 32.

Table 32

In-game behavioural scale items with corresponding feature mapping

Item	Mapped design feature		Item type
	Design feature dimension	**Design feature**	
What is the total number of icons and champions you own?	Expression	Avatar	Analytical
What is the total number of skins, wards, and chromas you own?	Expression	Design / Editing	Analytical
On average, how many times do you change an aspect of your profile every week? e.g., icon changes, border changes etc?	Expression	Profile	Reflective
What is the total number of draft pick games you have played until now?	Expression	Vote	Analytical
What is the total number of emotes you own?	Expression	Emotes	Analytical
On average, how many chests and capsule did you receive last month?	Improvement	Item Power-up	Reflective
What is the total number of riot points you own?	Improvement	Currency	Analytical
What is the total number of blue essence you own?	Improvement	Tokens	Analytical
On average, how many times do you take note or glance over your level progress bar when you play?	Improvement	Progress Bar	Reflective
What is your total mastery score?	Improvement	Points	Analytical
On average, how many hours do you spend on the client dashboard purposefully looking at your game history and stats every week?	Improvement	Dashboard	Reflective
What is your summoner level?	Competition	Rank	Analytical
On average, how many times do you check the leaderboards every week?	Competition	Leaderboard (Competition / Feedback)	Reflective
What is the overall number of games you have played (normal and ranked)?	Competition	PVP	Analytical

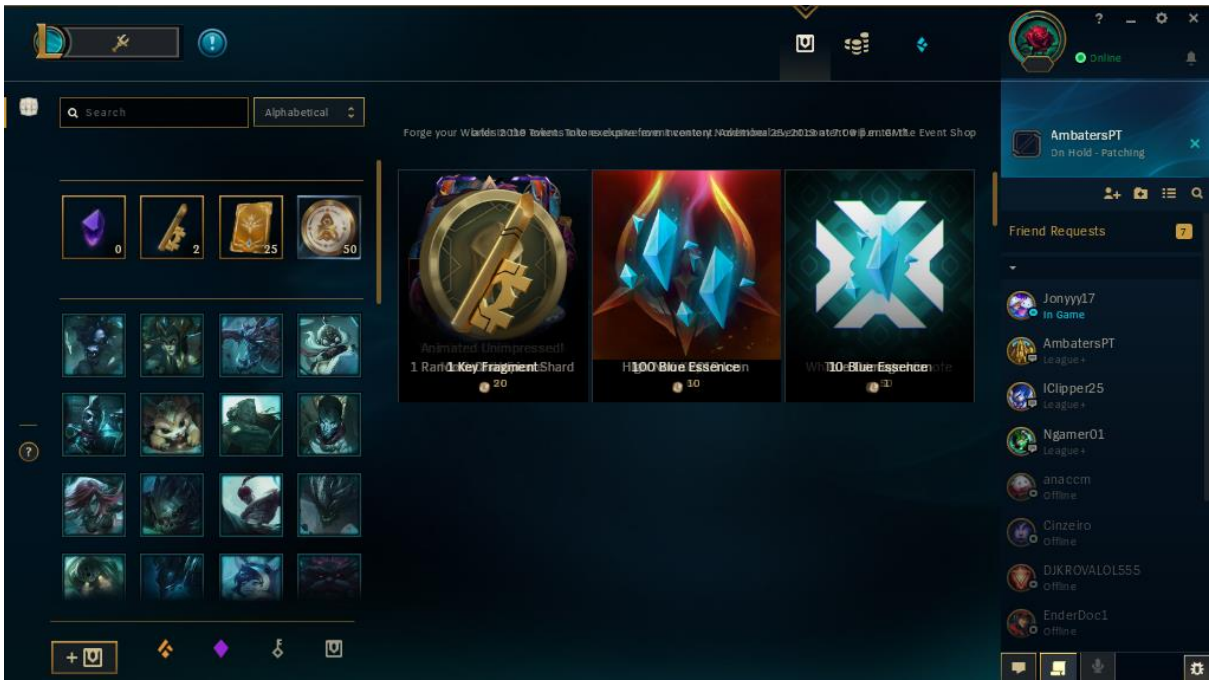
8.5.3.3 Demographics

Participants were asked to submit standard demographic data, such as age, ethnicity, gender, education status, country of domicile, employment status and marital status. In addition, general gaming related demographics were also collected, such as for how long participants have played video games (in years).

8.5.4 Procedure

Participants accessed the survey via a URL link provided in the recruitment advertisement, after which they were presented with the study information sheet. After providing consent, participants completed the demographics sheet, followed by the design feature rating task. Thereafter, participants were required to open the League of Legends player game client (Figure 33) so that they could complete the in-game behaviour scale. They were also advised to have a calculator to complete some of the questions. Upon completion of the in-game behaviour scale, participants were finally debriefed

Figure 33 – League of Legends player dashboard



8.5.5 Ethics approval

Ethical approval was granted by Birmingham City University's research ethics committee under the reference code: Lally/7996/sub2/R(A)/2021/Jan /BLSS FAEC

8.6 Results

The metric of user engagement for the present study was computed from in-game behavioural metrics from Riot Games', *League of Legends* (2009). Of the design features that could be mapped to engagement, only three design feature dimensions were identified: Expression, Improvement, and Competition. While preference for Expression and Improvement design features were found to significantly predict user engagement previous studies of this project, the in-game behavioural metrics available in the League of Legends player dashboard included several metrics that corresponded to Competition design features. A series of Pearson's correlation tests were conducted between design feature preference and user responses to the in-game behavioural metric scale, to identify how design feature preference would relate to variances in user engagement (operationalised as in-game behavioural metrics). Descriptive statistics showing which design features were mapped to in-game behavioural metrics, user engagement, as well as well as indications of normality (Skewness and Kurtosis), and the number of participants that provided responses across can be found in Table 36.

Table 36

Distribution statistics and mean scores for user engagement and design feature preference

Variable	Design feature mapping	Mean	SD	Skew (SE)	Kurt (SE)	*n*
User engagement (in-game behavioural metric)						
Expression						
Number of icons and champions owned	Avatar	200.29	104.72	0.36 (.472)	-1.20 (.918)	24
Number of skins, wards, and chromas owned	Design / Editing / Customisation	139.21	131.37	1.99 (.472)	5.64 (.918)	24
Weekly profile changes	Profile	2.48	10.15	4.86 (.472)	23.74 (.918)	24
Number of draft pick games played	Vote	2052.63	2407.32	2.02 (.472)	4.60 (.918)	24
Number of emotes owned	Emotes	60.08	40.47	1.24 (.472)	2.48 (.918)	24
Improvement						
Chests and capsule received	Item Power-up	8.71	7.55	2.16 (.472)	5.64 (.918)	24
Number of riot points owned	Currency	521.83	976.76	2.11 (.472)	3.28 (.918)	24
Number of blue essence owned	Tokens	15922.17	22518.79	1.96 (.472)	2.73 (.918)	24
Progress bar views	Progress Bar	7.79	21.29	3.98 (.472)	16.72 (.918)	24
Total mastery score	Points	274.13	163.28	0.09 (.472)	-1.08 (.918)	24
Time spent using dashboard	Dashboard	1.35	2.05	3.53 (.472)	14.71 (.918)	24
Competition						
Summoner level	Rank	202.33	94.72	0.06 (.472)	-1.21 (.918)	24
Weekly leaderboard check	Leaderboard (Competition / Feedback)	1.35	1.22	0.66 (.472)	-0.71 (.918)	24
Number of games played (normal and ranked)	PVP	2988.17	3058.57	1.54 (.472)	1.48 (.918)	24
Design features						

PVP	17.50	2.38	-0.40 (.472)	-0.81 (.918)	24
Leaderboard (Competition)	15.71	3.07	-0.86 (.472)	0.24 (.918)	24
Punish	10.83	5.22	-0.19 (.472)	-1.35 (.918)	24
Complementarity	13.63	5.31	-0.31 (.472)	-1.36 (.918)	24
Shared Goal	15.83	4.36	-1.21 (.472)	1.18 (.918)	24
Trade	10.75	5.58	0.14 (.472)	-1.40 (.918)	24
Friend Invite	18.58	2.65	-3.09 (.472)	11.18 (.918)	24
Text Chat	12.25	4.28	-0.08 (.472)	-0.54 (.918)	24
Voice Chat	15.42	4.36	-1.16 (.472)	1.10 (.918)	24
Emotes	12.21	4.88	-0.19 (.472)	-0.97 (.918)	24
Communal discovery	15.96	3.43	-0.36 (.472)	-1.03 (.918)	24
Trophy	14.17	4.72	-0.65 (.472)	-0.54 (.918)	24
Badges	13.75	5.08	-0.68 (.472)	-0.63 (.918)	24
Medal	13.42	5.13	-0.53 (.472)	-0.77 (.918)	24
Tokens	17.21	3.65	-1.98 (.472)	4.11 (.918)	24
Item Power-up	15.29	4.69	-0.89 (.472)	0.20 (.918)	24
Bonus	15.25	3.96	-0.77 (.472)	1.18 (.918)	24
Lottery	9.08	4.54	0.20 (.472)	-1.66 (.918)	24
Depletion	9.46	5.22	0.65 (.472)	-0.62 (.918)	24
Restriction	7.83	5.34	1.30 (.472)	0.27 (.918)	24
Demotion	12.00	5.24	-0.03 (.472)	-1.04 (.918)	24
Points	16.17	3.86	-1.61 (.472)	3.07 (.918)	24
Progress Bar	14.38	4.72	-0.94 (.472)	0.14 (.918)	24
Leaderboard (Feedback)	15.42	3.76	-0.83 (.472)	0.99 (.918)	24
Scarlett letter	7.13	3.60	1.20 (.472)	1.06 (.918)	24
Performance graphs	16.71	3.98	-2.13 (.472)	4.69 (.918)	24
Walkthrough	12.92	5.52	-0.48 (.472)	-1.27 (.918)	24
Tips / Hints	11.04	5.41	0.05 (.472)	-1.23 (.918)	24
Notification / Prompts	10.42	5.19	-0.06 (.472)	-1.49 (.918)	24
Cut Scenes	12.63	5.82	-0.31 (.472)	-1.48 (.918)	24
Storyline	14.75	4.15	-0.58 (.472)	0.14 (.918)	24
Currency	16.50	4.20	-1.54 (.472)	2.37 (.918)	24
Item Degradation	7.92	4.13	1.15 (.472)	1.56 (.918)	24
Dashboard	17.08	3.12	-1.23 (.472)	1.63 (.918)	24
Behavioural Momentum	17.21	4.77	-2.13 (.472)	3.77 (.918)	24
Levels	13.17	5.16	-0.38 (.472)	-1.06 (.918)	24
Barriers / Access	8.13	3.69	0.38 (.472)	-0.90 (.918)	24
Game Goal	14.33	5.24	-0.70 (.472)	-0.29 (.918)	24
Game Objective	16.63	4.59	-1.83 (.472)	3.30 (.918)	24
Boss Battles	16.67	4.37	-1.44 (.472)	1.60 (.918)	24
Beginners luck	12.17	4.98	-0.47 (.472)	-0.94 (.918)	24
Design / Editing / Customisation	14.58	4.71	-1.05 (.472)	0.32 (.918)	24
Decision Making	17.79	3.15	-1.41 (.472)	0.88 (.918)	24
Vote	13.83	4.72	-0.41 (.472)	-0.38 (.918)	24
Avatar	13.83	5.01	-0.76 (.472)	-0.20 (.918)	24
Profile	13.63	5.69	-0.50 (.472)	-1.06 (.918)	24

Rank / Status	17.21	4.14	-1.77 (.472)	2.65 (.918)	24

Notes: SD = Standard deviation, Skew = Skewness, Kurt = Kurtosis, SE = Standard error, n = Sample size.

Correlation scores and statistics between user engagement and design feature preference can be found in Table 33.

Table 33

Correlation summary of user engagement and design feature preference

Design feature dimension	**Design feature**	**Item / In-game metric**	***r(df)***	***r***	***p***
Expression	Avatar	Number of icons and champions owned	22	.-245	p > .050
Expression	Design / Editing / Customisation	Number of skins, wards, and chromas owned	22	.199	p > .050
Expression	Profile	Weekly profile changes	22	.263	p > .050
Expression	Vote	Number of draft pick games played	22	.-041	p > .050
Expression	Emotes	Number of emotes owned	22	.379	p > .050
Improvement	Item Power-up	Chests and capsule received	22	.-026	p > .050
Improvement	Currency	Number of riot points owned	22	.358	p > .050
Improvement	Tokens	Number of blue essence owned	22	.819	p > .050
Improvement	Progress Bar	Progress bar views	22	.024	p > .050
Improvement	Points	Total mastery score	22	.-061	p > .050
Improvement	Dashboard	Time spent using dashboard	22	.091	p > .050
Competition	Rank	Summoner level	22	.862	p > .050
Competition	Leaderboard (Competition / Feedback)	Weekly leaderboard check	22	.209	p > .050
Competition	PVP	Number of games played (normal and ranked)	22	.962	p > .050

8.6.1 Expression design features

8.6.1.1 Avatar

A Pearson's product-moment correlation analysis was conducted to assess the relationship between user preference for the Avatar design feature and total number of Icons and Champions users owned. The relationship between both variables was normally distributed, as assessed by a Shapiro-Wilk's test ($p > .05$), and there were no outliers. There was no statistically significant correlation between user preference for the Points design feature and total number of points earned, $r(22) = .-245$, $p > .050$, with preference for the Avatar design feature explaining 1% of the variation in total number of Icons and Champions owned.

8.6.1.2 Design / Editing

A Pearson's product-moment correlation analysis was conducted to assess the relationship between user preference for the Design / Editing design feature and the total number of skins, wards, and chromas each user owned. The relationship between both variables was normally distributed, as assessed by a Shapiro-Wilk's test ($p > .05$), and there were no outliers. There was no statistically significant correlation between user preference for the Design / Editing design feature and number of customisables users owned, $r(22) = .199$, $p > .050$, with preference for the Design / Editing design feature explaining 4% of the variation in number of customisables owned.

8.6.1.3 Profile

A Pearson's product-moment correlation analysis was conducted to assess the relationship between user preference for the Profile design feature and how frequently users reportedly changed an aspect of their profile. The relationship between both variables was normally distributed, as assessed by a Shapiro-Wilk's test ($p > .05$), and there were no outliers. There was no statistically significant correlation between user preference for the Profile design feature and how frequently users' would change an aspect of their profile, $r(22) = .263$, $p > .050$, with preference for the Profile design feature explaining 1% of the variation in how frequently users would change an aspect of their profile.

8.6.1.4 Vote

A Pearson's product-moment correlation analysis was conducted to assess the relationship between user preference for the Vote design feature and the total number of vote-related games the user has completed (draft pick games where a vote is necessary). The relationship between both variables was normally distributed, as assessed by a Shapiro-Wilk's test ($p > .05$), and there were no outliers. There was no statistically significant correlation between user preference for the Vote design feature and number of vote-related games completed, $r(22) = .-041$, $p > .050$, with preference for the Vote design feature explaining 1% of the variation in number of vote-related games completed.

8.6.1.5 Emotes

A Pearson's product-moment correlation analysis was conducted to assess the relationship between user preference for the Emotes design feature and the total number of character Emotes owned. The relationship between both variables was normally distributed, as assessed by a Shapiro-Wilk's test ($p > .05$), and there were no outliers. There was no statistically significant correlation between user preference for the Emotes design feature and number of character Emotes users owned, $r(22) = .379$, $p > .050$, with preference for the Emotes design feature explaining 4% of the variation in number of character Emotes owned.

8.6.2 Improvement design features

8.6.2.1 Item Power-up

A Pearson's product-moment correlation analysis was conducted to assess the relationship between user preference for the Item Power-up design feature and how many character enhancements users reportedly received in a given month (chests and capsules). The relationship between both variables was normally distributed, as assessed by a Shapiro-Wilk's test ($p > .05$), and there were no outliers. There was no statistically significant correlation between user preference for the Item Power-up design feature and monthly number of character enhancements received, $r(22) = .-026$, $p > .050$, with preference for the Item Power-up design feature explaining 1% of the variation in monthly character enhancements received.

8.6.2.2 Currency

A Pearson's product-moment correlation analysis was conducted to assess the relationship between user preference for the Currency design feature and how many spendable points (Riot Points) they possess. The relationship between both variables was normally distributed, as assessed by a Shapiro-Wilk's test ($p > .05$), and there were no outliers. There was no statistically significant correlation between user preference for the Currency design feature and total amount of spendable points they possessed, $r(22) = .196$, $p > .050$, with preference for the Currency design feature explaining 4% of the variation in total amount of spendable points the user possessed.

8.6.2.3 Tokens

A Pearson's product-moment correlation analysis was conducted to assess the relationship between user preference for the Tokens design feature and how much game tokens (Blue essence) they possess. The relationship between both variables was normally distributed, as assessed by a Shapiro-Wilk's test ($p > .05$), and there were no outliers. There was no statistically significant correlation between user preference for the Tokens design feature and total amount of game tokens they possessed, $r(22) = .-049$, $p > .050$, with preference for the Tokens design feature explaining 1% of the variation in total amount of game tokens the user possessed.

8.6.2.4 Progress Bar

A Pearson's product-moment correlation analysis was conducted to assess the relationship between user preference for the Progress Bar design feature and how frequently users reportedly would glance or pay attention to the Progress Bar when playing. The relationship between both variables was normally distributed, as assessed by a Shapiro-Wilk's test ($p > .05$), and there were no outliers. There was no statistically significant correlation between user preference for the Progress Bar design feature and frequency in viewing the Progress Bar when playing, $r(22) = .024$, $p > .050$, with preference for the Progress Bar design feature explaining 1% of the variation in frequency in viewing Progress Bar when playing.

8.6.2.5 Points

A Pearson's product-moment correlation analysis was conducted to assess the relationship between user preference for the Points design feature and total number of points earned (mastery score). The relationship between both variables was normally distributed, as assessed by a Shapiro-Wilk's test ($p > .05$), and there were no outliers. There was no statistically significant correlation between user preference for the Points design feature and total number of points earned, $r(22) = .-061$, $p > .050$, with preference for the Points design feature explaining 1% of the variation in total number of competitive matched played.

8.6.2.6 Dashboard

A Pearson's product-moment correlation analysis was conducted to assess the relationship between user preference for the Dashboard design feature and how many hours on average users' reportedly use the game dashboard to look at game history or stats per week. The relationship between both variables was normally distributed, as assessed by a Shapiro-Wilk's test ($p > .05$), and there were no outliers. There was no statistically significant correlation between user preference for the Dashboard design feature and average number of hours spent using the game dashboard, $r(22) = .091$, $p > .050$, with preference for the Dashboard design feature explaining 1% of the variation in average number of hours spent using the game dashboard.

8.6.3 Competition design features

8.6.3.1 Rank

A Pearson's product-moment correlation analysis was conducted to assess the relationship between user preference for the Rank design feature and the user's in-game Rank (summoner level). The relationship between both variables was normally distributed, as assessed by a Shapiro-Wilk's test ($p > .05$), and there were no outliers. There was no statistically significant correlation between user preference for the Rank design feature, and their in-game Rank, $r(22) = -.038$, $p > .050$, with preference for the Rank design feature explaining 1% of the variation in user in-game Rank.

8.6.3.2 Leaderboard

A Pearson's product-moment correlation analysis was conducted to assess the relationship between user preference for the Leaderboard design feature and how frequently users reportedly checked the game leaderboards. The relationship between both variables was normally distributed, as assessed by a Shapiro-Wilk's test ($p > .05$), and there were no outliers. There was no statistically significant correlation between user preference for the Leaderboard design feature and how often users' would check the game leaderboards, $r(22) = .209$, $p > .050$, with preference for the Leaderboard design feature explaining 4% of the variation in frequency of leaderboard checks.

8.6.3.3 PVP

A Pearson's product-moment correlation analysis was conducted to assess the relationship between user preference for the PVP design feature and total number of competitive matches they have played against other players. The relationship between both variables was normally distributed, as assessed by a Shapiro-Wilk's test ($p > .05$), and there were no outliers. There was no statistically significant correlation between user preference for the PVP design feature and total number of competitive matches they had played, $r(22) = .010$, $p > .050$, with preference for the PVP design feature explaining 1% of the variation in total number of competitive matched played.

8.7 Discussion

The primary aim of the present study was to use in-game behavioural metrics as a measurement of user engagement and assess the predictive value of the DFPS, by examining whether user design feature preference would correlate with reported in-game user behaviour. It was hypothesised that design feature preference would correspond to in-game user behaviour, such that increased preference for a given design feature would be associated with increases in the corresponding in-game behavioural metric (see Table 32). The results found no significant association between user engagement and design feature preference, therefore rejecting the hypothesis. Specifically, preference for Expression design features (Avatar; Design / Editing; Profile; Vote; and Emotes), Improvement design features (Item Power-up; Currency; Tokens; Progress Bar; Points; and Dashboard), and Competition design features (Rank; Leaderboard; and PVP) was not significantly associated with variances in user engagement.

Table 32

In-game behavioural scale items with corresponding feature mapping

Item	Mapped design feature		Item type
	Design feature dimension	Design feature	
What is the total number of icons and champions you own?	Expression	Avatar	Analytical
What is the total number of skins, wards, and chromas you own?	Expression	Design / Editing	Analytical
On average, how many times do you change an aspect of your profile every week? e.g., icon changes, border changes etc?	Expression	Profile	Reflective
What is the total number of draft pick games you have played until now?	Expression	Vote	Analytical
What is the total number of emotes you own?	Expression	Emotes	Analytical
On average, how many chests and capsule did you receive last month?	Improvement	Item Power-up	Reflective
What is the total number of riot points you own?	Improvement	Currency	Analytical
What is the total number of blue essence you own?	Improvement	Tokens	Analytical

On average, how many times do you take note or glance over your level progress bar when you play?	Improvement	Progress Bar	Reflective
What is your total mastery score?	Improvement	Points	Analytical
On average, how many hours do you spend on the client dashboard purposefully looking at your game history and stats every week?	Improvement	Dashboard	Reflective
What is your summoner level?	Competition	Rank	Analytical
On average, how many times do you check the leaderboards every week?	Competition	Leaderboard (Competition / Feedback)	Reflective
What is the overall number of games you have played (normal and ranked)?	Competition	PVP	Analytical

The most noticeable difference in the present study when compared with earlier studies of the PhD is the small sample size, which prevented the use of analytic methods that possessed more power (such as a Regression). According to Tabachnik and Fidell (2007), a sample ought to comprise of roughly 50 participants per predictor variable, which in the context of the present study would have required a total of 450 participants. Given the sample size was much smaller, a more powerful omnibus analysis could not be performed. One advantage of performing a regression would be to assess each preference rating as per an overall computed total.

At the core of this exists the issues with recruitment, which in future work can be improved. Initial expectations were such that because League of Legends is one of the most popular online games played worldwide, a large sample size could be acquired (thereby qualifying the data to be subject to more powerful statistical analyses). However, given the size of the League of Legends community, online social platforms (on which players interact) such as Reddit (r/LeagueOfLegends) are more heavily policed by moderators, less likely to endorse non-native projects, and are characterised by a higher frequency of new content being posted. One such example is some Reddit platforms not approving the recruitment of users for research if the project offers a form of participant inventive (such as a free game, as was done in previous studies of the PhD project). As such, in adhering to the rules of these Reddits, recruitment prospects were negatively impacted. Additionally, as the present study was from an external institution and non-native to the League of Legends community, moderators would also not be able to upvote or pin the thread to the homepage, which in previous studies of the project have helped increase rates of recruitment. Finally, due to the number of users of the League of Legends

Reddit, new content is posted frequently, meaning without successful endorsement by the moderator team (in the form of upvotes or pins), the recruitment advertisement reduced in visibility much quicker.

A solution to the issues encountered with recruitment and sample size could be to make use of server-side data in addition to client-side data, as opposed to relying exclusively on client-side data (information sourced and reported by the player). A limitation of using client-side data is that data could be misreported (due to reliance on the participant). In comparison, server-side data, (data that is automatically aggregated and extracted from the server), would resolve the limitation of misreported data and participant demand (to report) that could deter participation. A combination of both approaches could be most effective, with participant perceptions and preferences being provided by the participant, but their in-game behaviour being sourced from the server. One such example of this approach was seen in Kahn et al. (2015) who used both-client side and server-side data for League of Legends, with play motivations provided by the participants, while in-game behavioural data being sourced from the game servers for 18,627 players. In adopting this approach for future work however, two notable considerations must be made. Foremost, accessing server-side data would require collaboration with data handlers (e.g., Riot Games), which is likely to require more extensive permissions and compliance. Second, in the Kahn et al. (2015) example, scales which would record play motivations were presented to participants in-game, meaning they were able to complete the motivation scale in the League of Legends game menu and did not need to be externally directed. Not only did this ensure stronger recruitment prospects, but the pairing of participant motivation responses and in-game behavioural data was automatically achieved.

Beyond the issue of recruitment and sample size, non-significant correlations between user design feature preference and in-game behavioural data might also be explained by imprecise mappings between the DFPS and League of Legends in-game behaviour metrics (see Table 32). Given the lack of access to server-side data, the present study employed a design which required participants to access and report their available play statistics. As such, the present study was limited in how design features from the DFPS could be mapped onto in-game behavioural metrics, given the limited number of player statistics made available to the player.

To illustrate, the design feature of Design / Editing was mapped to the in-game behavioural metric of "total number of skins, wards, and chromas". Though skins, wards, and chromas do relate to the user being able to design or edit aspects of the game, the total number of these owned to not necessarily capture preference for the function of designing or editing. For instance, one could argue that total number of editable items a user owns is instead a greater indication of player skill, given that to purchase these items one needs in-game points, which are earned as a user level up. As such, there might not be a strong correlation between both data points (preference for the Design / Editing design feature and the in-game behavioural metric of total number of skins, wards, and chromas). A similar case of imprecise mappings could also be applied to the pairings between the DFPS Avatar design feature, and the total number of icons and champions owned.

Additionally, despite being associated with significant variance in user engagement, the present study was unable to include Accessibility design features in the measurement of design feature preference and corresponding user engagement. This was primarily due to the client-side data available for players of League of Legends in the player dashboard not having any metrics that corresponded to any of the Accessibility design feature, such as Tips / Hints, or Walkthroughs. Although these design features are present in the League of Legends video game, there was no way to ascertain how much time was spent interacting with them.

One method of improving the suitability of DFPS mappings with in-game behavioural metrics, would be to utilise server-side data. In addition to the previously highlighted benefits of reporting accuracy and participant load, server-side data is highly likely to provide several metrics which may not be accessible to the player but are more targeted and detailed. From a more focused selection of metrics, a more appropriate mapping of design features might be possible. Modern interactive systems, such as video games and mobile apps, pay close attention to the way in which they are used and interacted with. A variety of user activity is recorded, primarily to provide insights that can be monetised or leveraged to increase overall engagement (King et al., 2019). In the case of video games, Drachen et al. (2012) highlight some metrics which game designers can record, such as session lengths, points scored per minute, number of hits, number of misses, and time spent on a given page. In the case of the Design / Editing design feature, metrics

which assess how long a user spends or how often a user edits an aspect of the game would be a more appropriate mapping to correlate preference with.

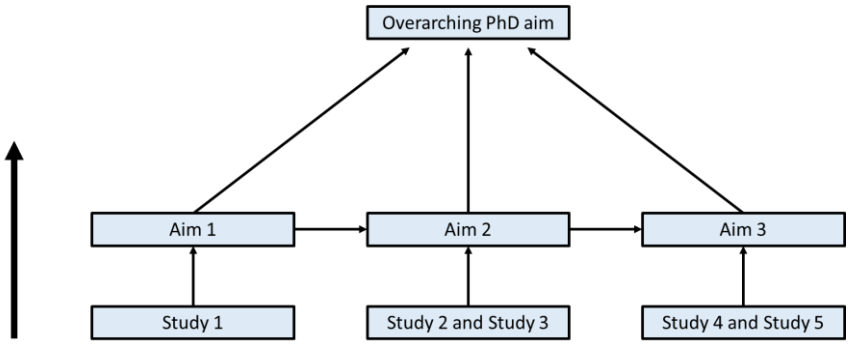
8.7.1 Summary

The primary outcome of the present study was to assess whether an alternative, more objective measurement of user engagement would validate previously identified relationships between user design feature preference and user engagement. The results of the present study indicated non-significant relationships; however, these findings were argued as being a product of methodological limitations, most of which were practical in nature. It was also suggested that working alongside games publishers who may provide telemetric data might benefit an in-game behaviour approach, though this is likely to require substantial cooperation from third parties.

9. General discussion

The overarching aim of the doctoral project was to identify how Adaptive Gamification could be more effectively designed when focusing on user needs, as indicated by individual differences. In fulfilling this overarching aim, three component research aims were conceptualised (Figure 3).

Figure 3 - Wireframe illustration of PhD research aims



9.1 Research aim fulfilment

9.1.1 Research aim one: Develop a measure to operationalise design feature preference

The first research aim was to develop a measure with which user design feature preference could be measured. Existing work in the area of Gamification was heavily characterised by a one-size-fits-all approach, such that user characteristics would receive little attention when gamifying a non-game context. In developing a measurement that could assess user design feature preference, researchers would be enabled to understand how Gamification can be more effectively adapted to meet user needs. The outcome of Study one was the development of the DFPS. The scale comprised of an original 37 design feature vignettes that describe the functionality and purpose of a given design feature.

Beyond its purpose, factors which arguably have prevented such a measure being developed in the past were also addressed, such as how design features are difficult to isolate. The use of vignettes also provides the benefit of being able to exclusively focus on one design feature, such that preference reported can be more reliably interpreted as being related to a single design feature. Further, Study one also addressed the issue of terminological consistency in existing Gamification literature, as the DFPS comprises of

vignettes that detail the function of a given design feature, thus there is clarity on what the participant is reporting a preference for. The issue of specificity was also addressed by the DFPS, given that each vignette would comprise of a similar number of words, and would describe the function of a given design feature, there would be a high level of specificity provided, such that researchers could easily determine what the participant was providing a preference rating for. The aspect of contextual suitability was also addressed, with the DFPS comprising only of design features that could be applied in both real-life and digital gamified contexts (e.g., a Demotion design feature, which sees a user's rank or status reduce following a failure of some sort). Finally, the broader issue of limited design feature scope was also resolved by the DFPS, with it comprising of over 37 design features (increased to 47 in Study two, Table 16, section 5.5.3.1).

The contribution of initial DFPS model development could be summarised in two main ways; the first, is the novel method of using vignettes to conceptualise and describe design features, for which preference and other variables can easily be compared with. Not only do vignettes provide a cost-effective and easily manipulatable research measurement (Veloski et al., 2005), but they also mitigate a variety of other limitations that emerge in Gamification literature (see previous paragraph). Moreover, the DFPS model is structured in way which allows for additional design features to be added in future work (given the criteria of including a new design feature is for that feature to be isolatable, can be described textually, and will achieve high levels of readability). Thus, the DFPS can be updated in line with advancements in gaming or technology (and new design features which may emerge).

Secondly, the initial DFPS was further developed after being used in conjunction with measures of individual differences, revealing how user design feature preference could be predicted by user needs and characteristics (Study two and three). The initial development of the DFPS also enabled the exploration of alternative measurements of user engagement (Study four and five) Collectively, the findings from these chapters provide insight on how Adaptive Gamification can be more effectively designed and researched, with design building on the insights on user needs and characteristics, and research building on insights from measurements of user engagement.

Taken further, subsequent model development occurred in Study two, in the form of an EFA. Given the total of 47 design features, assessing how preference for each design

feature would relate to variances in an individual difference would not be statistically possible without substantially increasing the risk of a Type I error (due to multiple testing) (Andrade, 2019). As such, an EFA was selected as the most effective variable reducing technique that could explain the underlying correlations between each design feature. Moreover, the EFA was also expected to return a more parsimonious model of design feature preference (Raykov & Marcouilides, 1999), which was desirable given the number of variables (design features) involved. The EFA generated a nine-factor solution (Figure 6, section 5.6.1).

In Study three, a CFA was conducted; a statistical technique that seeks to verify the factor structure of a set of variables (Suhr, 2006) which in the context of DFPS model development would be to confirm and verify the nine-factor solution generated by the EFA in Study two. The suitability of conducing a CFA comprised of four main arguments; it is more appropriately used in the later stages of model development, most suited when using a new sample not previously used, more suited when used with a data set of more than 200 participants, and if all original loadings of the preceding EFA score above .3. The CFA confirmed the overarching factor structure of the EFA, however, there were some differences in design feature loading (Table 23, section 6.7.1). The final model provides a confirmed summary of inter-relatedness between design features. One benefit of this is the theoretical guide on how broader design feature dimensions (such as Expression or Competition) might be implemented with constituent design features (e.g., Avatar or Design / Editing). Moreover, in seeking to implement broader design feature dimensions in Gamification design, the DFPS provides multiple options on how this could be achieved, of which researchers could opt to implement a combination of constituent design features, or all of them together.

9.1.2 Research aim two: Identify how the individual differences of users could predict design feature preference

The second research aim was to further validate the DFPS by examining whether variances in user personality, motivation, and reported Gamefulness could predict design feature preference (and subsequent user engagement). As the causal role of dimension preference on user engagement was only observed across the EIA dimensions. As such, all subsequent analyses focused on these dimensions out of the original nine, however it

must be noted that the focus on EIA dimensions was underpinned by the outcomes of Study two, which used a sample comprising mostly of WoW players.

Table 24

Finalised loading of design feature dimensions that predicted variance in user engagement

Accessibility	**Expression**	**Improvement**
Tips / Hints	Avatar	Item Power-up
Being able to receive tips and hints when playing	Being able to represent yourself via a virtual model/sprite/signature	Being able to receive beneficial items / power ups after completing challenges or tasks
Walkthrough	Design / Editing / Customisation	Currency
Having the option to receive a step-by-step guide on how to complete tasks or play the game	The option to edit or design aspects of the game (e.g., avatar, environment, inventory)	Being able to spend your in-game money/currency on game content
Beginners Luck	Profile	Tokens
Help in achieving a high rate of success when completing the first few tasks	Being able to immediately convey several aspects of yourself to other players	Being able to earn tokens after completing challenges, that can be used to buy game content
Notifications	Vote	Progress Bar
Being able to receive key notifications and updates when playing	Having the opportunity to vote on something (e.g., map, weapon, rules)	Being able to see how close you are to reaching a milestone in a bar format
	Emotes	Dashboard
	Being able to express your emotion and feelings through your avatar behaviour (such as jumping or dancing)	Being able to access game information, such as your game history, resources, profile, friends list, achievements etc
		Points
		Being able to access game information, such as your game history, resources, profile, friends list, achievements etc
		Bonus
		The scenario in which an unexpected or additional reward is received

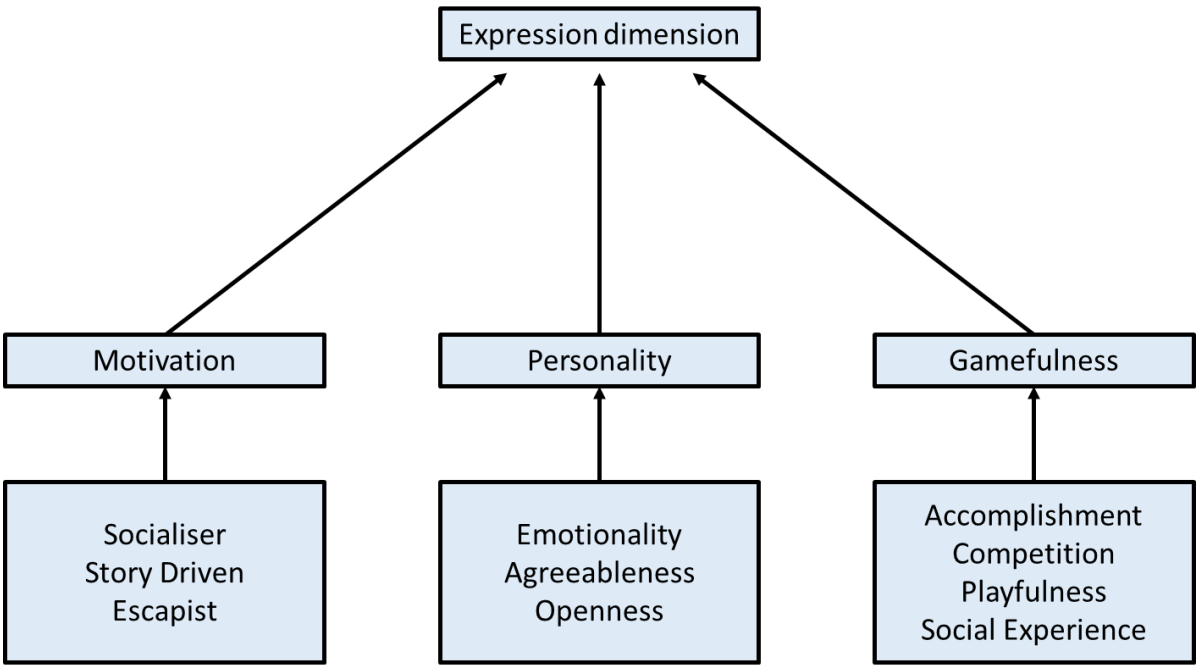
Across all analyses conducted in Study two and three, a range of convergences between variances in user motivations, personality traits, and values of Gamefulness had emerged. Study two examined how design feature preference would predict variance in motivation and personality, while Study three examined how design feature preference would predict variance in reported Gamefulness. Taken together, although they cannot be interpreted as entirely comprehensive of player preference, given that preference is

generally dynamic and evolves or change overtime (Heifetz et al., 2007), there existed several commonalities between design feature preference and motivation, personality, and Gamefulness. These commonalities point towards latent trends in user disposition and preference. While design feature preference was measured by the DFPS, user motivation was measured using the Trojan Player Typology (Kahn et al., 2015), user personality was measured using the HEXACO personality scale (Ashton & Lee, 2004), and user Gamefulness was measured using the Gameful Experience Questionnaire (Högberg et al., 2019).

9.1.2.1 Expression design features

The individual differences which significantly predicted variance in user preference for Expression design features can be found in Figure 34. Within the context of Expression design features, variance in preference that emerged across all three individual difference measures highlight a set of user needs for social proximity to others, an appreciation of knowledge, and the use of fantasy as a mechanism for escapism. These user needs are discussed in relation to how Expression design features possess strong potential at fulfilling these needs.

Figure 34 - Individual differences associated with preference for Expression design features



9.1.2.1.1 Social experience

Users who foster the motivation of Socialiser tend to play games to build and maintain social relationships, frequently play with others, and enlist or encourage others to join them (Kahn et al., 2015). These user tendencies complement the trait tendencies of Agreeableness, such as being highly cooperative and collaborative (Ashton and Lee, 2004) and the Gamefulness dimension of Social-experience, which is to value the experiences that emanate from the direct or indirect presence of other users (Högberg et al., 2019). In the context of design feature preference, user needs for social proximity can be effectively fulfilled by Expression design features, given that most of the expression design features function within the more social aspects of gaming (Pringle, 2015). For example, one function of an Avatar design feature and Profile design feature can be viewed as front facing representations of the user that other users are exposed to, which research indicates is a cornerstone to the development of online societies wherein users regularly socialise and communicate with one another (Novak, 2022). Furthermore, the Design / Editing design feature is likely to encourage social interactions, with research highlighting that the game character's outward attractiveness can determine their social status as well as interpersonal attraction from other users (Lo, 2008). The role of a Design / Editing design feature may enable users to make their online character more attractive, thereby leading to more social interactions, as research indicates that attractive online character is more likely to acquire higher in-game social status and subsequent interest from other users in comparison to less attractive online characters (Lo, 2008). In addition, many elements of a game that users can edit or design can often be exchanged, discussed, or traded with other users. For example, the collection, trading, and purchasing of rare items in World of Warcraft and RuneScape), which prompt a range of different purchasing considerations to be made and discussed with other users, such as whether the item is being fairly traded, it's perceived value, If it can be customised, if it will bring enjoyment, and the perceived social status of those participating in the transaction (Guo & Barnes, 2012). Moreover, the Vote design feature inherently provides the opportunity to join other users in decision making processes and cooperation, while the Emotes design feature provides the opportunity for users to convey emotion to a more interactive degree (Kim et al., 2022). As such, user need for social proximity consists with increased preference for Expression design features.

9.1.2.1.2 Appreciation for knowledge

Users who foster the Story-driven motivation tend to desire interesting stories in the gaming world, and to learn about the backgrounds of game characters and underlying lore (Kahn et al., 2015). These user tendencies complement the trait tendencies of Openness to Experience, for which individuals with high levels are intellectually curious, seek absorption in knowledge, and take interest in the unusual or unconventional (Ashton & Lee, 2004). In the context of video games, research indicates that users who are higher in Openness to Experience tend to orient their game interests to exploration and understanding more of the game world. For instance, Worth and Book (2014) report that Openness to Experience was associated with users having completed more exploration related achievements in a MMORPG (during with world discovery occurred).

In the context of preferring Expression design features, the Design / Editing design feature could likely fulfil the user need of knowledge, given that the user is indirectly exposed to the microdetail of a game world or game characters during the process of editing or customising those same microdetails. Moreover, as the Design / Editing design feature is often operated with an in-game inventory (from which users can select what they wish to use to replace the element they are customising), the user is exposed to a greater degree of "surrounding" game world knowledge (which usually accompanies descriptions of a given item that are contained within the players' inventory). Research examining how the modern video games engage players in lore exposition also confirms this, as designers aim to convey the game world and lore via microdetails of a video game thereby avoiding overwhelming the user (Wolek, 2022). One such example of microdetail conveyance is character design, which is thought to convey key concepts of a world, the languages used, and colour theory. As such, the user need for knowledge consists with increased preference for Expression design features.

9.1.2.1.3 Fantasy as a mechanism for escapism

Users who foster the Escapist motivation tend to use games to escape from real-life and fulfil real-life deficits by using the digital world and fantasy as the primary mechanism (Kahn et al., 2015). The aversive position towards real-life characteristic of users who foster Escapist motives could be classed as a similarity to the aversive disposition of those who score high in the Emotionality personality trait, described as individuals who

experience anxiety in response to life stresses (Di Blasi et al., 2019). As such, aversive attitudes towards real-life is a shared position. An increased preference for Expression design features was also correlated with increased user value for the Gamefulness dimension of Playfulness, which describes user pursuits of experiences that are driven by imagination and free from rules (Högberg et al., 2019). In the context of preferring Expression design features, the aversive attitudes towards real-life that are characteristic of the Escapist motivation and the Emotionality personality trait closely complement the Gamefulness value of Playfulness, and collectively these user needs are fulfilled effectively by Expression design features.

For example, the Avatar design feature can provide users with an opportunity to create an artificial and digital model of themselves (thereby engaging in fantasy) that enable projections which remedy perceived personal deficits (e.g., user appearance or persona) (Mazlan & Bakar, 2013). Moreover, the Design / Editing design feature targets the user need to operate with little restriction of rules, given that the function of this design feature is to customise and edit features of the game world. The reinvention of oneself that is possible within the digital world could also extend beyond how the user creates their online image by also influencing the way in which they interact with others (Moore et al., 2017). For instance, users can adopt different emotions in the digital fantasy context by dissociating with their identity in real-life. One such example of this is the disinhibition effect, which in the context of video games sees users acting and communicating in a manner substantially different to how they would interact with others in a real-life context (Beres et al., 2021). As such, the use of fantasy as a mechanism is a collective of users across motivation, personality, and Gamefulness, and consists with an increased preference for the Expression design feature dimension.

9.1.2.1.4 Summary

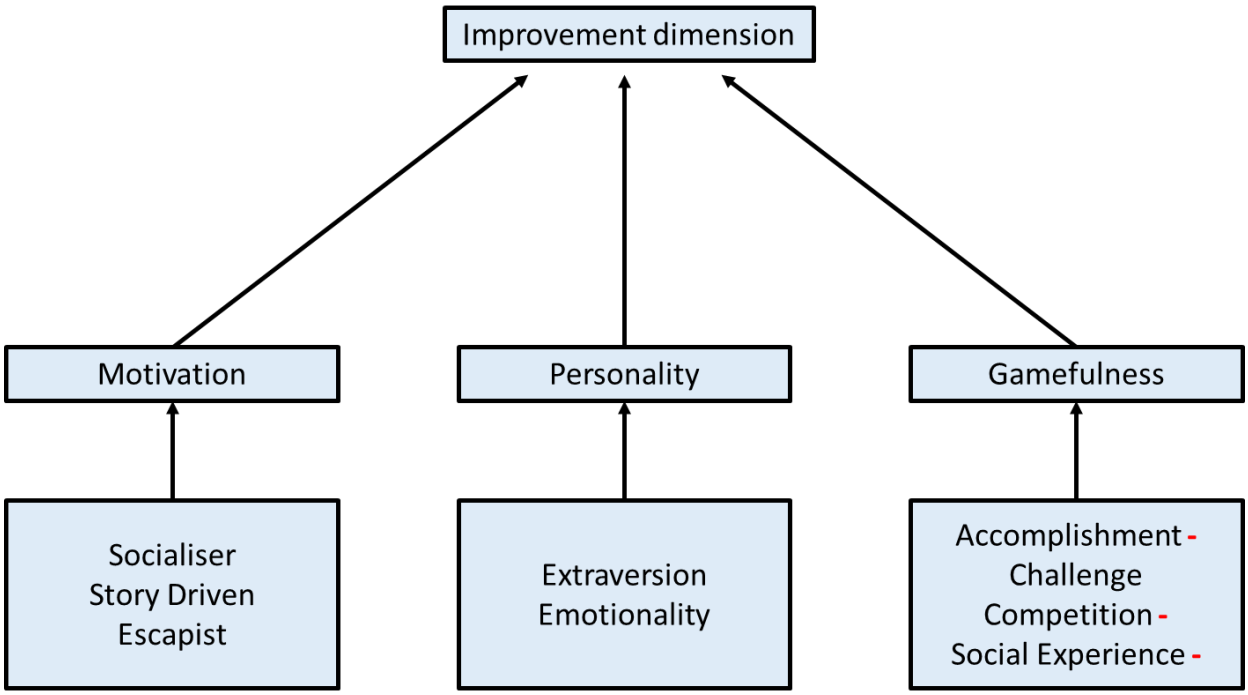
In summary, variance in user preference for Expression design features was predicted by variance in user motivation, personality, and Gamefulness. Although distinct, there existed some convergence across these areas of individual differences, indicating a broader set of user needs that Expression design features potentially fulfil, such as the need for a social experience, appreciation for knowledge, and using fantasy as a mechanism for escapism. Considering these findings, the fulfilment of these user needs

may be effectively achieved by the implementation of Expression design features, the primary implication of which being a potential increase in overall user engagement.

9.1.2.2 Improvement design features

The individual differences which significantly predicted variance in user preference for Improvement design features can be found in Figure 35. Within the context of Improvement design features, variance in preference that emerged across all three individual difference measures highlight a set of user needs for social proximity to others and the use of fantasy as a mechanism for escapism, like that which emerged for preference of Expression design features. These user needs are discussed in relation to how Improvement design features possess strong potential at fulfilling these needs.

Figure 35 - Individual differences associated with preference for Improvement design features



9.1.2.2.1 Social experience

Similar with the convergences associated with preference for Expression design features, increased preference for Improvement design features was also associated with a similar need for social proximity, demonstrated in the commonality between the Socialiser motivation and the Extraversion personality trait. Defined as the tendency to feel confident about oneself and enjoy social interactions, as well as often adopt leadership

positions, the motivation and personality trait pairing could be explained by the design feature for which preference varied. Compared with the personality dimension of Agreeableness, the Extraversion trait is likely to be more applicable to users who reported increased preference for Improvement design features, given that they are likely to be more mindful and attentive to the development of their competencies and skills. Specifically, research indicates a positive feedback loop between Extraversion and the development of skills, with Extraverts being likely to possess effective skills in social capacities, and the development of these skills are likely to lead to higher level of Extraversion (Matthews & Gilliland, 1999). In the context of preferring Improvement design features, the tendency to improve skill is well facilitated, particularly with the Progress Bar design feature (that indicates how close a user is to achieving a milestone), the Item Power-up design feature (which increases the user skill), and the Tokens design feature (which enables the user to enhance their skill level). Moreover, given that individuals with higher levels of Extraversion are more likely to pursue social interactions confidently, it might also be argued that these users would make use of Improvement design features to augment their confidence in the online social environment. Demonstrating the role of higher competence and more social interaction in online environments, Ducheneaut et al. (2006) report that high level users (who possess a higher rank than others) generally have a higher status in MMORPGs, which corresponds to a greater level of social influence and subsequently a higher number of social interactions (Lo, 2008).

Building on the notion that high Extraversion is characterised by an increased desire to self-improve, users who reported increased preference for Improvement design features were also associated with an increase in value for the Gamefulness dimension of Challenge. Defined as the need to experience demand such that user ability and skill is tested, the value placed on experience of Challenge coincides with the Extravert tendency to pursue self-improvement, given that improvements are likely to emerge following a challenging experience. In support, research by Bentea and Anghelache (2012) highlights that Extraversion is predictive of the pursuit of competitive challenges, with those who are higher in Extraversion demonstrating an increased level of performance compared with lower levels of Extraversion. In the context of Improvement design feature preference, the Points, Progress Bar, and Item Power-up design features are particularly

suited. Collectively, the user need for social proximity consists with increased preference for Improvement design features.

9.1.2.2.2 Fantasy as a mechanism for escapism

Similar to the patten of results found with Expression design features, a need for fantasy as a mechanism for escapism was also found to be associated with Improvement design features. Given the tendency for those motivated by Escapism seek to disassociate from real-life to fulfil deficits (Kahn et al., 2015), and that users who report greater levels of Emotionality tend to adopt a highly aversive position towards the stresses of real-life (Högberg et al., 2019), an increased preference of these users for Improvement design features can be expected, as the function of Improvement design features is that of enabling the user to improve in-game potential and competence.

In the context of escaping the real world, Improvement design features might also be of significant preference to these users. Research indicates that reaching a state of flow and immersion often corresponds with high level of engagement and focus when playing a game (Faas et al., 2014). As Improvement design features focus on providing the user with the means to improve their in-game skills and/or character, it is likely that users who want to escape from real-life will engage considerably with these design features. For example, the Item Power-up design feature (which enhances user skill following receipt of an in-game item) as well as the Currency or Tokens design feature (both of which enable the augmentation of a character depending on what is purchased) may prompt users to direct a significant level of focus on completing in-game tasks, such that they can improve their skills, and indirectly encourage immersion in the game world (and therefore escape from the real world). Supporting literature highlights that in-game challenge, which is a precursor to achieving a state of immersion, is often a strong predictor of an increase in overall engagement (Hamari et al., 2016). In the context of these findings, users may spend more time playing in order to upskill so that they can overcome game challenges, a result of which would be to improve the potential for immersion and effectively “escape” from real life.

9.1.2.2.3 Summary

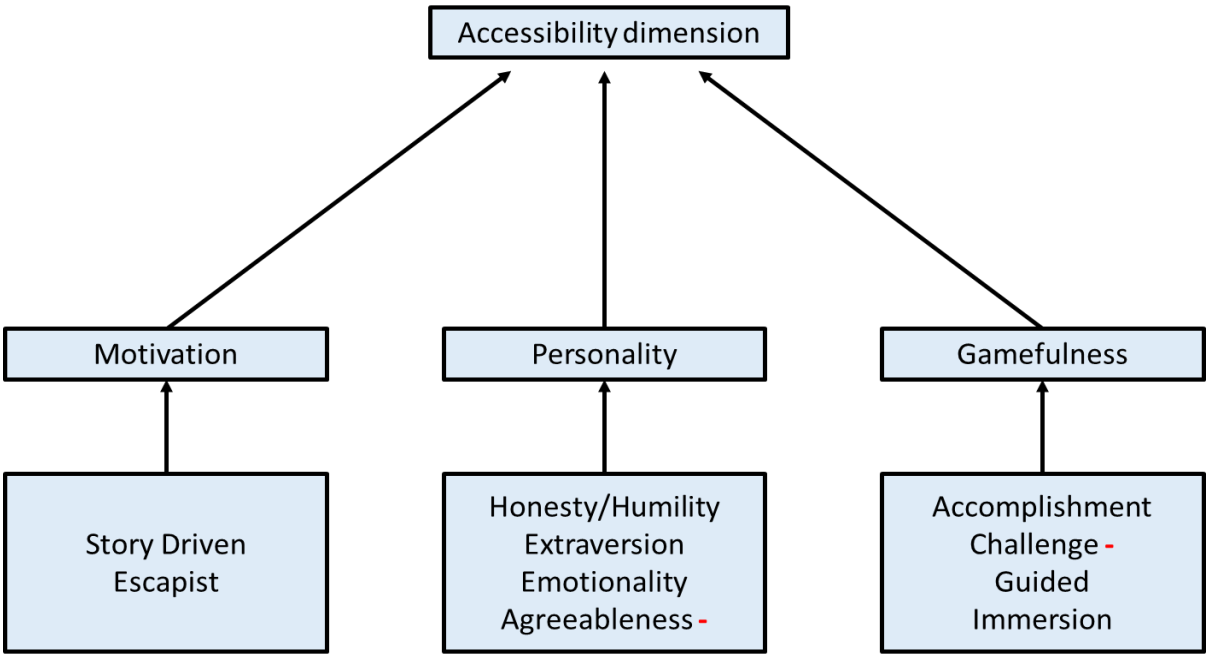
In summary, variance in user preference for Improvement design features was predicted by variance in user motivation, personality, and Gamefulness. Although distinct, there

existed some convergence across these areas of individual differences, similar to patterns which emerged in the analyses for preference for Expression design features, such as the user need for social experience and using fantasy as a mechanism for escapism. Considering these findings, the fulfilment of these user needs may be effectively achieved by the implementation of Improvement design features, the primary implication of which being a potential increase in overall user engagement.

9.1.2.3 Accessibility design features

The individual differences which significantly predicted variance in user preference for Accessibility design features can be found in Figure 36. Within the context of Improvement design features, the collective emergence of each individual difference facet highlights a set of user needs for world immersion and using fantasy as a mechanism for escapism. These user needs are discussed in relation to how Accessibility design features possess strong potential at fulfilling these needs.

Figure 36 - Individual differences associated with preference for Improvement design features



9.1.2.3.1 World immersion

Users who foster the Story-driven motivation tend to heavily focus on the gaming world narrative and story, as well as the stories of game characters (Kahn et al., 2015). What

can be argued as a lack of primary interest in features of the game that involve other players (e.g., competition) could also be true of users who report high levels of the Honesty-Humility personality trait. For example, those who score high in Honesty-Humility, tend to avoid manipulation of others for personal gain (Ashton & Lee, 2004). Manipulating others for personal gain is often a feature of modern games that encourage competition, especially those which encourage 1 vs 1 competition (wherein misleading or deceiving an opponent is often a requisite for victory). Supporting research from Ueno et al. (2017) details the presence of dark triad trait tendencies in competitive sports (Machiavellianism, Psychopathy, and Narcissism), highlighting the inherent adversarial dynamic of competition can prompt interpersonal behaviours that conflict with behaviours otherwise expected of those who score in the Honesty-Humility personality trait. In the context of video games, one example of a game that more explicitly promotes behaviours that conflict with a Honest-Humility trait disposition, is the popular mobile game Among Us (2018), which encourages users to deceive and manipulate their team to win.

Additionally, high Honesty-Humility is characterised by a lack of interest in wealth, luxuries, and elevated social status. (Ashton & Lee, 2004). In support, a review conducted by Ashton et al. (2014) further indicate that those who score high in Honesty-Humility are far less motivated by monetary gain or social status, which in the context of games could be approximated to a lack of interest in the social and monetary reward systems intrinsic of most modern games. For example, earning in-game currency, items, and rewards, and player rank or level. Taken together, those who score high in the Story-driven motivation and high in the Honesty-Humility personality trait, are likely to orient their design feature preference to those which do not directly or indirectly relate to other users.

Building on the user need for world immersion, a third convergence exists in the value for the Guided Gamefulness dimension, which is described as valuing the experience of being guided on how and what to do in the game world (Högberg et al., 2019). Users who are interested in acquiring a deeper understanding of the game world, would also highly value the Guided Gamefulness dimensions. For example, the Tips / Hints design feature and the Walkthrough design feature provide users with information and guidance on how they can orient and learn the functions of the game, such that they are able to operate

effectively. Existing literature also substantiates the relationship between guidance and Honesty-Humility, with those who score higher in this trait also tending to offer guidance more readily to others as well as more readily seek and accept guidance from others (Fang et al., 2019). As the individual differences associated with increased preference for Accessibility design features indicate a preference for individual play along with a greater interest in the game world, it would be consistent to also expect these users to prefer design features that enable the exploration of the game world without involving users.

9.1.2.3.2 Fantasy as a mechanism for escapism

Along with high levels of the Escapist motivation (Kahn et al., 2015) and Emotionality personality trait (Ashton & Lee, 2004), a third convergence found in association with preference for Accessibility design features, was the increased value for the Immersion Gamefulness dimension, which is described as valuing the experience of being absorbed in the game world and experiencing a sense disassociation from the world (Högberg et al., 2019).

In the context of Accessibility design features, the Tips / Hints and Walkthrough design features serve the function of onboarding a user to how they can effectively operate in the game-world. Given that achieving a state of Immersion requires a balance between user competence and game difficulty (Sigailov-Lanfranchi, 2019), it could be argued that to become immersed, a familiarity with game controls and processes is a prerequisite. As such, increased preference for Accessibility design features complements a user need for fantasy as a mechanism of escapism in both an experiential capacity (engaging with the game story while avoiding engagement with other users) as well as an in an operational capacity (learning how to operate in the game to achieve states of immersion, thereby escaping real-life).

9.1.2.3.2 Summary

In summary, variance in user preference for Accessibility design features was predicted by variance in user motivation, personality, and Gamefulness. Although distinct, there existed some convergence across these areas of individual differences, similar to patterns which emerged in the analyses for preference for Expression and Improvement design features, such as the user need for using fantasy as a mechanism for escapism, as well as the new user need for world immersion. Considering these findings, the fulfilment of

these user needs may be effectively achieved by the implementation of Accessibility design features, the primary implication of which being a potential increase in overall user engagement.

9.1.2.4 Summary of individual differences and design feature preference

Given the varied areas of individual differences examined (motivation vs personality vs Gamefulness), similarities that emerged provide researchers a range of positions from which Gamification design can be more effectively adapted to increase user receptivity and engagement. For instance, the fine tuning of Gamification design to better facilitate competitive play could be approached from a position of user motivation, user personality, or both. Moreover, the findings from Study two and three provide researchers with an initial guide on how aspects of user motivation, personality, and Gamefulness are related, such that future work can further examine how they may be combined and better leveraged to improve Gamification design, and consequently improve overall user engagement.

9.1.3 Research aim three: Substantiate whether identified relationships would correspond to objective differences in user engagement

As part of further model development, the third research aim was to substantiate whether the findings from Study two; that user preference for EIA design features was associated with significant changes in user engagement, would also occur for more objective measurements of user engagement.

9.1.3.1 Study four

Study four examined how preference for design feature dimensions would correspond to user engagement as measured by real-time user behaviour. Real-time user behaviour was measured by assessing how long users spent playing a sequence of online browser games, each of which were selected based on the design features they primarily comprised of. By measuring the users demonstrated activity, the methodological issues associated with participant recollection and self-report would be mitigated, and the validity of the DFPS further substantiated by accurately predicting real-time user behaviour. From the DFPS, only five dimensions were examined in Study four; Expression, Improvement, Competition, Difficulty, and Loss. The exclusion of the Accessibility, Reward, Cooperation,

and Narrative design feature dimensions were due to difficulties in identifying games which exclusively or predominantly comprised of these design features limitations. For instance, Accessibility design features can often be found towards the start of a game (such that the user is successfully onboarded and familiarised with the game). Given the use of online browser games, there were no games which predominantly comprised of Accessibility design features, nor was there the ability to restrict participants to the part of a given online browser game wherein Accessibility design features are used i.e., at the start.

The primary outcome of Study four found that of all relationships examined, none were found to be significant. As such, user design feature preference did not predict variance in real-time user behaviour. Despite these findings, Study four does highlight a series of methodological considerations which future work and the wider field of Gamification would benefit in making. The contribution of Study four suggest that online browser games might not be an effective method at curating gaming behaviour that users are likely to exhibit when playing games in a non-experimental environment. For instance, browser games are inherently simpler when compared to console or PC alternatives, given they do are hosted online and not stored locally. Given the role of demand in prompting engagement (Forkosh &Drake, 2017), the use of online browser games, which are intrinsically less likely to require a level of cognitive demand that is equivalent to what is required by more sophisticated games (that are locally stored and can be operated with more computing power), it could be argued that online browser games are likely to be less engaging. If participants are used to gaming experiences that are more demanding in general, then browser games would be less engaging, and therefore less likely to emulate the conditions of normal games, wherein natural play behaviour could emerge. In the context of Gamification and the wider field, this outcome can guide methods employed by researchers to curate natural play behaviour, such that less complex games might not be an effective solution. In support, instances where Gamification has overused a smaller number of design features, such that it is less elaborate and complex, has often been ineffective (Nacke & Deterding, 2017).

However, the difference in engagement elicited by users when playing browser games compared with what would be expected of them when playing more sophisticated console or PC games could be mediated by user characteristics. For instance, if

participants tend to spend more time playing browser games to begin with, then the issue of graphic difference and subsequent reduction in engagement, is unlikely to persist. As such, another contribution of Study four is to highlight a wider need to focus on user characteristics in Adaptive Gamification research.

The findings from Study four did not consist with the wider findings of the project, such that preference for EIA design features did not correspond to differences in user engagement. Beyond methodological limitations, another perspective could be that the relationships observed in earlier studies may not exist. A key feature of earlier studies in this project was that user engagement was retrospectively estimated and then reported by participants. Building upon those findings relies on the assumption that participants accurately reported their play behaviour, however, this cannot be ascertained. In using a more direct measure of user engagement, the relationship between preference and engagement did not persist.

Theoretical arguments might also support the notion that preference for EIA design feature do not relate to differences in engagement. For instance, the concept of diminished reward sensitivity proposes that via repeat exposure, the incentive to acquire rewards reduces over time (Capa & Bouquet, 2018). In the context of Improvement design features, such as Points, Bonus, or Item Power ups, are arguably more ubiquitous across games, therefore as users are more likely to repeatedly receive rewards via these design features as play duration increase, the potential for diminished reward sensitivity increases. As such, increases in preference may not correspond to reported increases in user engagement, given that interaction with the design features do not elicit the same behavioural response.

Other examples of where design feature preference might not impact user engagement might also be explained by in-game paywalls, which can be described as a monetisation strategy whereby users are prevented from accessing parts, features, or areas of the game unless payments are made (Gallo et al., 2016). In the context of Expression design features, such as the Avatar, and Design / Editing, the availability of customisable items might not be readily available to players, such that payment is required to acquire different items that can be used to customise and express oneself. For instance, the MMORPG Guild Wars 2 (2012) is one of many examples wherein users can spend real money to purchase in-game cosmetics and items (Ommen, 2018). As such, although users

may report preference for Expression design features, corresponding variance in engagement may not emerge, given that they are unable to interact with the given design features (and expected responses will not occur).

9.1.3.2 Study five

Study five examined how design feature preference would correspond to user engagement when measured by in-game metrics, which were reported by users after consulting an in-game dashboard that automatically records player behaviours. In measuring user engagement by in-game behaviour and testing whether previously identified relationships between preference and engagement would persist, the DFPS would be further validated as being effective at predicting changes in player in-game behaviour. In-game behavioural metrics were operationalised in the form of mapping a given in-game metric to a given design feature (see Table 32). For example, the number of icons or champions a user owned would be correlated with the preference the user expressed for an Avatar design feature. From the DFPS, only three dimensions were examined in Study five; Expression, Improvement, and Competition. The exclusion of the Accessibility, Reward, Cooperation, Narrative, Difficulty, and Loss dimensions were due to difficulties in mapping design features to a given in-game behavioural metric. For instance, from the Accessibility dimension, there existed no in-game metrics that would provide insight on how often or how many Tips and Hints the user was presented with, or the time spent going through a Walkthrough.

The primary outcome of Study five found that of all relationships examined, none were found to be significant. As such, user design feature preference did not predict variance in user engagement (i.e., in-game behaviour). Despite these findings, however, Study five does highlight a series of methodological considerations with which future work and the wider field of Gamification would benefit in making. For example, the use of in-game behavioural metrics as a measurement of engagement might need more effective mapping.

The use of in-game metrics in gaming research, specifically to provide more quantified indications of user activity and behaviour is well supported. For example, Kiili et al. (2018) report that in-game metrics were highly effective at predicting learning gains made by students after playing a mathematical learning game. A noticeable difference

between Study five and this example, concerns the origin of the in-game metrics that were used to collect data. In the maths game, in-game metrics were purposefully designed to measure a variable of interest. For instance, a metric such as game performance was calculated by looking at the number of correctly solved estimation tasks, comparison tasks, and ordering tasks.

In the context of Study five, however, in-game metrics were adapted and retrospectively assigned, meaning a degree of inaccuracy would be more likely to emerge when assessing whether a design feature would map effectively onto an existing LoL in-game metric. To illustrate, the design feature of Design / Editing was mapped to the in-game behavioural metric of "total number of skins, wards, and chromes", which although do relate to the user being able to design or edit aspects of the game, may not capture the entire scope of a user's design and editing activity. Moreover, mapping a design feature to precise in-game metrics may exclude other aspects of a game that could be more representative at demonstrating a user's tendency in that given area. Therefore, a contribution of Study five to the wider area of Adaptive Gamification, is to highlight that mapping design features to in-game metrics could be made more effective if they are purpose built and highly adaptable to the design feature of interest. Doing so would provide a more representative account of how design feature preference corresponds to in-game player behaviour, thereby providing a more reliable indication of whether reported user design feature preference will correspond to actual player behaviour.

It might be argued, however, that this method is costly and not time effective. The production of a small mobile game requires a development team, testing, software licensing, and significant finance in place (up to $150,000) (Starloop Studios, 2022). One method which may enable researchers to circumvent the need to have a purpose-built game from which in-game metrics can be measured, could be to acquire in-game behavioural data from game servers and character profiles. For instance, Kahn et al. (2015) made use of server-side player data to substantiate whether motivational tendencies reported were consistent with natural play activity. Similarly, Yee et al. (2012) used a range of in-game metrics to substantiate user reported motivations when playing a MMORPG, by using a data scraping methodology which would track player behaviour over time by assessing player profile updates and changes. In the context of Study five, though behavioural data was collected by examining player profiles, this was collected

using player self-report (and not server-side data or data scraping, both of which are automated and require no involvement from the user themselves), which required the user to consult their own player statistics and report them. This approach was likely to be more cost-effective and not require cooperation of third parties (such as game publishers), however, was limited in how reliable the reported statistics were. Despite this, Study five findings contribute to the wider area of Adaptive Gamification by providing an example of how in-game metrics could be operationalised within research if server-side data access is not available. The findings and discussions also highlight a range of methodological improvements that researchers could build on if employing this approach in future work.

9.2 Implications

The findings from this doctoral thesis have several implications. First and foremost, the development of the DFPS provides researchers within the area of Adaptive Gamification a cost effective and validated research instrument which measures user design feature preference. At the core of the Adaptive Gamification paradigm exists the principle that tailoring the design of Gamification to meet the needs of the end user is the most effective method at increasing user engagement. With the DFPS, researchers are enabled to identify which design features are most preferred by users (and therefore which design features can be prioritised within Gamification design to increases the potential of fulfilling user needs). The psychometric and conceptual design of the DFPS also improves on existing limitations within the Gamification area, such that design feature isolation can be achieved and subsequent areas of association (such as how preference may relate to user characteristics) can be researched. Beyond approaching Adaptive Gamification from a more granular perspective, the DFPS also introduced a comprehensive and broader range of defined and operationalised design features (a total of 47) that previous models did not (Arnab et al., 2015; da Rocha Seixas et al., 2016; Sailer et al., 2017). As the DFPS comprises of design feature vignettes, researchers are also enabled to make future additions as well as adjusting or editing the vignettes to fit specific research contexts or aims. For example, in contextualising design feature vignettes to represent a design feature in the context of a given game title or genre to enhance the participant familiarity or comprehension.

Further, the findings of the doctoral project also contribute to the growing understanding of how Gamification can target user needs, specifically by establishing how design feature preference is predicted by user psychology. While this project focused on concepts of motivation, personality, and Gamefulness, the application of the DFPS is not limited to these constructs. For example, gaming research indicates that user intelligence can have an impact on overall play tendencies and preferences (Dang et al., 2019). The DFPS therefore could also be measured alongside user intelligence, such that any latent relationship between intelligence and design feature preference could further inform how Gamification can be designed to meet a different type of user needs more effectively. Other constructs which could also be measured alongside the DFPS include age (Greenber et al., 2010), gender (Rehbein et al., 2016), native language (Tekofsky et al., 2016), and socioeconomic status (Thorn et al., 2014). Thus, the insights and relationships that can be found in using the DFPS, as well as the versatility and cost-effectiveness in applying or editing the DFPS, place it as highly versatile and inexpensive research instrument that can make significant contribution to the field of Adaptive Gamification.

Beyond its applicability, the findings generated in using the DFPS in relation to user characteristics can also be progressed in future work. For instance, cross cultural differences in personality and motivation are well documented (Kaushal & Kwantes, 2006; Morling & Lee, 2017), therefore assessing whether identified relationships in this doctoral project persist in different cultures could reveal whether styles of Adaptive Gamification might be better suited to different users depending on their prevailing or native culture. For instance, research indicates that motives to play games can lead to different outcomes depending on culture, with playing to fulfil an escapist motivation being associated with negative outcomes for western players (e.g., loneliness, social withdrawal, and time wastage), but leading to positive outcomes for eastern players (e.g., social connection, recreation, and stress-coping) (Hussain et al., 2021). Though differences in user characteristics due to culture has not yet been researched in the context of Gamefulness, understanding how there may exist variance in what aspects of a gaming experience users value, depending on their culture, will also inform Adaptive Gamification strategy. Related work can also be found in research relating to genre preference and culture, which supports the perspective that overarching preferences can be mediated by culture (Barza & Memari, 2014).

The project also introduced a range of different measurement of user engagement, as well as how they can be operationalised for research and associated strengths and weaknesses that may be improved in future work. Earlier studies in the project utilised self-report measures, though participant data could not be validated due to the subjectivity associated with self-report. Subsequent studies utilised more objective measurements of user engagement that offered more validation, focusing on task-performance, and in-game user behaviour. Collectively, the use of these measures provides empirical examples of how all three types of engagement measures could be implemented into research. For instance, in the case of task-performance, a contribution made by Study four demonstrates the use of free-to-play online browser games, which can be emulated in future work focusing on how users engage with a given game based on its design feature profile. Furthermore, free-to-play online browser games also provide an environment in which multiple players interact with one another. For example, one of the five games used in Study four, Bullet Force, is a competitive shooting game where large teams can play against each other. Given the highly competitive and cooperative nature of this online game, research which might be interested in examining online player discourse (Ensslin, 2015) or online player interactions (Cook et al., 2019) could make use of similar free-to-play online browser games as a natural online game environment from which observations and data could be collected.

In the case of in-game behavioural metrics, Study five contributes to the wider field of gaming research by providing an example of how in-game metrics can be identified and mapped to variables of interest from a front-end system position (i.e., what is freely available). Although back-end system data is likely to provide more objectivity, and a greater range of behavioural metrics that are continuously aggregated (Kahn et al., 2015), a benefit of using the method demonstrated in Study five is the circumvention of needing game publisher involvement with research projects, which would otherwise be costly. In a broader sense, using this method would also ensure that the use of in-game metrics in research could potentially increase, given the barrier of insufficient project funding could be surmounted. Furthermore, by highlighting that in-game metrics can be populated by front-end player data that is freely available, a potentially wider range of questions related to user behaviour could be answered. For instance, along with metrics demonstrated in Study five, other games, such as SMITE, Halo, and DOTA 2, all provide users with a range of player statistics (usually via an API). As each game differs in the

experience it affords the user, a wider range of freely available player data can be operationalised to explain difference in user activity and behaviour.

The effectiveness of sampling from Reddit was also demonstrated across the first four studies in the project. The benefits of using Reddit include being able to find participants who possess key sample characteristics of interest (e.g., players), as well as the relatively simple process of acquiring cooperation from the moderators of a given Reddit. Although each Reddit is governed by a set of rules relative to the given community, many do not restrict recruiting, and only require prior approval from moderators to continue (usually consisting of providing study information, details of what the participants must do, and in some cases providing a trial or preview of the survey). Researchers conducting future online work may benefit from approaching Reddit, especially if studies require participants that play video games.

9.2.1 Implications for the workplace

Beyond contributions to research, the relationships identified throughout this project can also be leveraged to improve behavioural outputs across a range of contexts. For instance, Gamification has been used extensively in occupational contexts, wherein conventional systems and processes are gamified to increase employee engagement and improve performance (Ferreira-Oliveira et al., 2017). The outcomes of this project provide indications on how Gamification could be tailored to users based on design feature preference, as well as user motivation, personality, and perceived Gamefulness. In the context of the workplace, a diverse set of motivations and personalities can often emerge (Sackett & Walmsley, 2014), therefore, approaching the implementation of design features from an Adaptive Gamification perspective and tailoring the design features employees interact with in a gamified system to complement their personal characteristics could improve the overall impact of Gamification.

In cases where effective Gamification is realised within the workplace, a range of benefits may emerge for both employees and broader business structures. For example, in the case of the user, effective Gamification would lead to user needs being more effectively fulfilled, and an improvement in workplace performance. A cascading outcome of improved performance can see employees achieve more positive affective states, such as feeling happy and accomplished (Azmat & Iriberri, 2010). In such cases, the broader

business can also benefit, with improved productivity often leading to more realised commercial and business value and mitigate sunken costs (El-Telbany & Elraga, 2017).

One example of how the PhD findings could directly be leveraged into a workplace scenario, would be to enable forms of expression. A considerable body of research indicates that when employees perceive that their ability to express themselves is inhibited, organisational dissent and job dissatisfaction can often occur (Kassing, 2000). The Expression factor of the DFPS provides a range of design features which may enable employees to express themselves, thereby potentially protecting organisations from limiting employee expression. For instance, research indicates that the use of emojis can influence perceptions of employees (Riordan, 2020), which in the context of the DFPS is similar to that of the Emotes design feature. The implementation of the Emotes design feature may provide employees a method at expressing opinions on work-related matters (such as jumping for joy to represent excitement or happiness, or the shaking of the head to represent disagreement), thereby preventing dissatisfaction.

9.2.2 Implications for student learning

A similar environment in which adapting Gamification to meet user needs could yield highly positive impact is with student learning. Several studies have gamified learning to improve student performance (Sailer & Homner, 2020), however, often these attempts are not successful due to a lack of user engagement (Roy & Zaman, 2017). Research indicates that learning can be heavily influenced by personality and motivation (Kim et al., 2013; Shepard et al., 2018). The project findings of how individual differences relate to engagement and design feature preference could inform and guide how the Gamification of learning contexts might be adapted to meet user needs more effectively (and improve overall performance prospects).

For example, conventional gamified learning systems often seek to improve learning by encouraging student competition, such as when applying a points feature or a leaderboard (Kiryakova et al., 2014). However, student receptivity to competition is not universal, with contrasting research indicating that learning may be improved within a context of cooperation, and not competition Boud et al. (1999). For instance, research on peer learning, which sees students consolidate their learning by sharing, teaching, and exchanging their understanding on a given topic with peers, indicates that students can

often benefit from this type of learning as it reduces the stress and anxiety that ensues following a mistake (Wessel, 2015).

One example of how the PhD findings could directly be leveraged into student learning, would be to support students to engage in peer learning. The Cooperation factor of the DFPS provides a range of design features that could enable students to learn with one another, thereby realising the benefits of peer learning. For instance, the Complementarity design feature, which mandates that an objective can only be fulfilled with the help of another player, or the Communal Discovery design feature which enables tasks to be completed as a community.

9.2.3 Implications for health and fitness

A third context within which the findings from this project could be applied is within the area of health and fitness. Across the world, several social initiatives and efforts have been started to combat a global obesity epidemic, which is estimated to have increased in total cost since from 2015 to over £5bn in the UK alone (Lobstein, 2015). The role of Gamification in supporting users to lose weight and lead a healthier lifestyle is well supported. The Gamification of health and fitness, specifically in a mobile app capacity, is extensive (Cotton & Patel, 2018). A popular example being *MyFitnessPal*, a mobile app which supports users in their health, eating, and fitness tracking (Wolf & Weiger, 2018). Engaging with health and fitness apps may improve user health prospects and weight management, given that they assist in the management of food consumption and exercise, both of which are critical components to maintaining a healthy lifestyle (Dunn et al., 2011). Additionally, research highlights the role of individual differences at influencing the way in which a user approaches fitness, as well as their likelihood of engaging in fitness-related activities (Smith & Blumstein, 2008). Building on the Adaptive Gamification approach, mobile apps which seek to support health and fitness may be more engaging to users if they are to be adapted to meet and fulfil the differing user needs.

One example of how the PhD findings could directly be leveraged into a health and fitness scenario, would be to support forms of improvement. A fundamental behavioural tendency associated with successful weight loss is an awareness of progress (Ingels et al., 2017) characterised by the individual maintaining an understanding of their objectives and the required level of work needed to achieve them. The Improvement factor of the

DFPS provides several design features which can support users in their awareness of goals as well as the level of work outstanding to meet those milestones. For instance, the Progress Bar design feature could assist users in situating the progress they have made, and the progress they need to still make, which aligns with existing findings on the power of self-regulation in supporting weight loss (Gokee-LaRose et al., 2009).

9.3 Thesis conclusion

Overall, this doctoral project aimed to enable researchers with a validated and operational measure with which user design feature preference could be understood in relation to user needs. Central to this aim was the development of the DFPS and its effectiveness at measuring user design feature preference, as well as applying the DFPS to further understand how design feature preference could be predicted by user characteristics, such as personality, motivation, and Gamefulness. The use of the DFPS and the findings from each study demonstrate its use case, and provides insight into which design features users prefer, and subsequently how Gamification can be more effectively adapted. Findings from Study one informed how items of the DFPS could be most effectively generated, while findings from Study two and Study three refined the factor structure of the model, and highlighted how motivation, personality, and Gamefulness could predict user design feature preference. Moreover, these studies also highlighted how user engagement could be improved in Gamification, by identifying how user preference, as collected by the DFPS, could predict variance in overall engagement, however it must also be acknowledged that such findings drew from samples which predominantly comprised of WoW players. Findings from the final two studies also revealed the methodological considerations researchers must make when determining how user engagement is to be effectively measured. Study four explored a measurement of user engagement wherein user activity could be recorded and measured during play activity, though it was concluded that the implementation of this form of measurement would have benefited if other methodological considerations (such as the artificial nature of requiring participants to successively play multiple games) were made. Study five sought to improve on this form of measurement by deriving engagement data from in-game player metrics that would be recorded over time, therefore representing more natural play behaviours. In conclusion, future work can build upon several of the outcomes from this doctoral project in several capacities. For example, utilising the DFPS

to explore how other individual difference dimensions may inform user design feature preference. Finally, future work can gain insight into how more objective measurements of user engagement can be more effectively operationalised in research.

10. Bibliography

Aarseth, E., Bean, D. A., Boonen, H., Colder Carras, M., Coulson, M., Das, D. van Rooij, A. (2016). Scholars' open debate paper on the World Health Organization ICD-11 Gaming Disorder proposal. *Journal of Behavioral Addictions*, *6*, 1-4. https://doi.org/10.1556/2006.5.2016.088

Abuhamdeh, S., Csikszentmihalyi, M., & Jalal, B. (2015). Enjoying the possibility of defeat: Outcome uncertainty, suspense, and intrinsic motivation. *Motivation and Emotion*, *39*(1), 1-10. https://doi.org/10.1007/s11031-014-9425-2

Allam, A. M. (2017). Investigating the Relationship Between Gamers' Personality Traits and their Preferred Video Game Genres.

Andersen, E., O'Rourke, E., Liu, Y.-E., Snider, R., Lowdermilk, J., Truong, D., . . . Popovic, Z. (2012). *The impact of tutorials on games of varying complexity* Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, Texas, USA. https://doi.org/10.1145/2207676.2207687

Andrade, C. (2019). Multiple Testing and Protection Against a Type I (False Positive) Error Using the Bonferroni and Hochberg Corrections. *Indian Journal of Psychological Medicine*, *41*(1), 99-100. https://doi.org/10.4103/IJPSYM.IJPSYM_499_18

Anglim, J., & O'connor, P. (2019). Measurement and research using the Big Five, HEXACO, and narrow traits: A primer for researchers and practitioners. *Australian Journal of Psychology*, *71*(1), 16-25. https://doi.org/10.1111/ajpy.12202

Arai, S., Sakamoto, K., Washizaki, H., & Fukazawa, Y. (2014). *A Gamified Tool for Motivating Developers to Remove Warnings of Bug Pattern Tools*. https://doi.org/10.1109/IWESEP.2014.17

Arnab, S., Lim, T., Carvalho, M. B., Bellotti, F., de Freitas, S., Louchart, S., . . . De Gloria, A. (2015). Mapping learning and game mechanics for serious games analysis [https://doi.org/10.1111/bjet.12113]. *British Journal of Educational Technology*, *46*(2), 391-411. https://doi.org/https://doi.org/10.1111/bjet.12113

Ashton, M. C., & Lee, K. (2001). A theoretical basis for the major dimensions of personality.

European Journal of Personality, *15*(5), 327-353. https://doi.org/10.1002/per.417

Ashton, M. C., & Lee, K. (2009). The HEXACO–60: A Short Measure of the Major Dimensions of Personality. *Journal of Personality Assessment*, *91*(4), 340-345. https://doi.org/10.1080/00223890902935878

Association, E. S. (2015). 2015: Sales, Demographic and Usage Data - essential Facts about the Computer and Video Game Industry. In.

Bailey, E. R., Matz, S. C., Youyou, W., & Iyengar, S. S. (2020). Authentic self-expression on social media is associated with greater subjective well-being. *Nature Communications*, *11*(1), 4889. https://doi.org/10.1038/s41467-020-18539-w

Bang, M., Gustafsson, A., & Katzeff, C. (2007, 2007//). Promoting New Patterns in Household Energy Consumption with Pervasive Learning Games. Persuasive Technology, Berlin, Heidelberg.

Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2017). Studying student differentiation in gamified education: A long-term study. *Computers in Human Behavior*, *71*, 550-585. https://doi.org/https://doi.org/10.1016/j.chb.2016.08.049

Barbel, K., Robert, F. B., Daniel, H. H., & Herzog, A. R. (1997). Question Difficulty and Respondents' Cognitive Ability: The Impact on Data Quality. *Journal of Official Statistics*, *13*(2).

Barr, M., & Copeland-Stewart, A. (2021). Playing Video Games During the COVID-19 Pandemic and Effects on Players' Well-Being. *Games and Culture*, *17*(1), 122-139. https://doi.org/10.1177/15554120211017036

Bartle, R. (1996). Hearts, clubs, diamonds, spades: Players who suit MUDs. *Journal of MUD research*, *1*(1), 19.

Barza, S., & Memari, M. (2014). Movie Genre Preference and Culture. *Procedia - Social and Behavioral Sciences*, *98*, 363-368. https://doi.org/10.1016/j.sbspro.2014.03.427

Basak, C., Voss, M. W., Erickson, K. I., Boot, W. R., & Kramer, A. F. (2011). Regional differences in brain volume predict the acquisition of skill in a complex real-time strategy videogame. *Brain and Cognition*, *76*(3), 407-414. https://doi.org/https://doi.org/10.1016/j.bandc.2011.03.017

Ben-Artzi, E., & Hamburger, Y. A. (2001). Private Self-Consciousness Subscales: Correlates with Neuroticism, Extraversion, and Self-Discrepancy. *Imagination, Cognition and Personality*, *21*(1), 21-31. https://doi.org/10.2190/LL2E-5RFH-45FJ-HXL7

Bergstrom, K. (2021). Anti-social social gaming: community conflict in a Facebook game.

Critical Studies in Media Communication, *38*(1), 61-74. https://doi.org/10.1080/15295036.2020.1863441

Bessière, K., Seay, A., & Kiesler, S. (2007). The Ideal Elf: Identity Exploration in World of Warcraft. *Cyberpsychology & behavior : the impact of the Internet, multimedia and virtual reality on behavior and society*, *10*, 530-535. https://doi.org/10.1089/cpb.2007.9994

Billieux, J., Van der Linden, M., Achab, S., Khazaal, Y., Paraskevopoulos, L., Zullino, D., & Thorens, G. (2013). Why do you play World of Warcraft? An in-depth exploration of self-reported motivations to play online and in-game behaviours in the virtual world of Azeroth. *Computers in Human Behavior*, *29*(1), 103-109. https://doi.org/https://doi.org/10.1016/j.chb.2012.07.021

Birk, M., Friehs, M., & Mandryk, R. (2017). *Age-Based Preferences and Player Experience: A Crowdsourced Cross-sectional Study*. https://doi.org/10.1145/3116595.3116608

Birk, M., Mandryk, R., & Atkins, C. (2016). *The Motivational Push of Games: The Interplay of Intrinsic Motivation and External Rewards in Games for Training*. https://doi.org/10.1145/2967934.2968091

Blocker, K., Wright, T., & Boot, W. (2014). Gaming Preferences of Aging Generations. *Gerontechnology*, *12*, 174-184. https://doi.org/10.4017/gt.2014.12.3.008.00

Bogen, K. (1996). *The Effect of Questionnaire Length on Response Rates: A Review of the Literature*. U.S. Bureau of the Census.

Bowman, S. (2018). Immersion and Shared Imagination in Role-Playing Games. In (pp. 379-394). https://doi.org/10.4324/9781315637532-22

Boyle, G. J. (2008). Critique of the five-factor model of personality. In (Vol. 1). https://doi.org/10.4135/9781849200462.n14

Braakmann, N. (2009). The Role of Psychological Traits for the Gender Gap in Full-Time Employment and Wages: Evidence from Germany. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1357346

Braun, B., Stopfer, J. M., Müller, K. W., Beutel, M. E., & Egloff, B. (2016). Personality and video gaming: Comparing regular gamers, non-gamers, and gaming addicts and differentiating between game genres. *Computers in Human Behavior*, *55*, 406-412. https://doi.org/https://doi.org/10.1016/j.chb.2015.09.041

Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., & Pidruzny, J. N. (2009). The development of the Game Engagement Questionnaire: A measure of

engagement in video game-playing. *Journal of Experimental Social Psychology*, *45*(4), 624-634. https://doi.org/https://doi.org/10.1016/j.jesp.2009.02.016

Brown, T. A. (2003). Confirmatory factor analysis of the Penn State Worry Questionnaire: Multiple factors or method effects? *Behaviour Research and Therapy*, *41*(12), 1411-1426. https://doi.org/https://doi.org/10.1016/S0005-7967(03)00059-7

Brunborg, G., Mentzoni, R., & Frøyland, L. R. (2014). Is video gaming, or video game addiction, associated with depression, academic achievement, heavy episodic drinking or conduct problems? *Journal of Behavioral Addictions*, *3*, 27-32. https://doi.org/10.1556/JBA.3.2014.002

Butler, Y. G., Someya, Y., & Fukuhara, E. (2014). Online games for young learners' foreign language learning. *ELT Journal*, *68*(3), 265-275. https://doi.org/10.1093/elt/ccu008

Bédard, A., Garcia-Aymerich, J., Sanchez, M., Le Moual, N., Clavel-Chapelon, F., Boutron-Ruault, M.-C., . . . Varraso, R. (2015). Confirmatory Factor Analysis Compared with Principal Component Analysis to Derive Dietary Patterns: A Longitudinal Study in Adult Women. *The Journal of Nutrition*, *145*(7), 1559-1568. https://doi.org/10.3945/jn.114.204479

Böckle, M., Novak, J., & Bick, M. (2017). *TOWARDS ADAPTIVE GAMIFICATION: A SYNTHESIS OF CURRENT DEVELOPMENTS*.

Cade, R., & Gates, J. (2016). Gamers and Video Game Culture: An Introduction for Counselors. *The Family Journal*, *25*(1), 70-75. https://doi.org/10.1177/1066480716679809

Campbell, L., Simpson, J. A., Stewart, M., & Manning, J. (2003). Putting Personality in Social Context: Extraversion, Emergent Leadership, and the Availability of Rewards. *Personality and Social Psychology Bulletin*, *29*(12), 1547-1559. https://doi.org/10.1177/0146167203256920

Capa, R. L., & Bouquet, C. A. (2018). Individual differences in reward sensitivity modulate the distinctive effects of conscious and unconscious rewards on executive performance. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00148

Caplan, S. E. (2003). *Preference for online social interaction: A theory of problematic Internet use and psychosocial well-being* [doi:10.1177/0093650203257842].

Caselli, G., & Spada, M. M. (2015). Desire thinking: What is it and what drives it? In

Castel, A., Pratt, J., & Drummond, E. (2005). The effects of action video game experience on the time course of inhibition of return and the efficiency of visual search. *Acta psychologica*, *119*, 217-230. https://doi.org/10.1016/j.actpsy.2005.02.004

Cattell, R. B. (1966). The Scree Test For The Number Of Factors. *Multivariate Behavioral Research*, *1*(2), 245-276. https://doi.org/10.1207/s15327906mbr0102_10

Chapman, B., Duberstein, P., Sörensen, S., & Lyness, J. (2007). Gender differences in Five Factor Model personality traits in an elderly cohort. *Personality and Individual Differences*, *43*, 1594-1603. https://doi.org/10.1016/j.paid.2007.04.028

Chappel, C. (2020, December 2020). PUBG Mobile Tops List of Billion-Dollar Mobile Games in 2020.

Chen, J., He, S., & Yang, X. (2020). Perceived Fairness and User Engagement: A Quasi-Natural Experiment on Impacts of Rule Breach and Recovery Measures in Online Gaming. *Available at SSRN 3666536*.

Clifford, S., Iyengar, V., Cabeza, R., & Sinnott-Armstrong, W. (2015). Moral foundations vignettes: a standardized stimulus database of scenarios based on moral foundations theory. *Behavior Research Methods*, *47*(4), 1178-1198. https://doi.org/10.3758/s13428-014-0551-2

Codish, D., & Ravid, G. (2014). Personality based Gamification: How different personalities percive Gamification. *ECIS 2014 Proceedings - 22nd European Conference on Information Systems*.

Collins, M. (2007). Engaging Self-Actualisation through Occupational Intelligence. *Journal of Occupational Science*, *14*(2), 92-99. https://doi.org/10.1080/14427591.2007.9686589

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis, 2nd ed*. Lawrence Erlbaum Associates, Inc.

Cook, C., Conijn, R., Schaafsma, J., & Antheunis, M. (2019). For Whom the Gamer Trolls: A Study of Trolling Interactions in the Online Gaming Context. *Journal of Computer-Mediated Communication*. https://doi.org/10.1093/jcmc/zmz014

Cook, R. D. (1977). Detection of Influential Observation in Linear Regression. *Technometrics*, *19*(1), 15-18. https://doi.org/10.2307/1268249

Cotton, V., & Patel, M. S. (2018). Gamification Use and Design in Popular Health and Fitness Mobile Applications. *American Journal of Health Promotion*, *33*(3), 448-451. https://doi.org/10.1177/0890117118790394

Craipeau, S., & Seys, B. (2005). Jeux et Internet : quelques enjeux psychologiques et sociaux. *Psychotropes*, *11*. https://doi.org/10.3917/psyt.112.0101

Cruz, C., Hanus, M., & Fox, J. (2015). The need to achieve: Players' perceptions and uses of extrinsic meta-game reward systems for video game consoles. *Computers in Human Behavior*, *71*. https://doi.org/10.1016/j.chb.2015.08.017

da Rocha Seixas, L., Gomes, A. S., & de Melo Filho, I. J. (2016). *Effectiveness of Gamification in the engagement of students*. In [doi:10.1016/j.chb.2015.11.021]. Elsevier Science.

Dang, L., Zhang, M., Leong, K., & Wu, A. (2019). The Predictive Value of Emotional Intelligence for Internet Gaming Disorder: A 1-Year Longitudinal Study. *International Journal of Environmental Research and Public Health*, *16*, 2762. https://doi.org/10.3390/ijerph16152762

Daughenbaugh, R., Daughenbaugh, L., Surry, D., & Islam, M. (2002a). Personality Type and Online versus In-Class Course Satisfaction. *Educause Quarterly*, *3*.

Daughenbaugh, R., Daughenbaugh, L., Surry, D., & Islam, M. (2002b). Personality Type and Online versus In-Class Course Satisfaction. *Educause Quarterly*, *25*(3), 71-72.

David, F. N. (1998). *Games, gods, and gambling: A history of probability and statistical ideas*. Courier Corporation.

de Freitas, S., Gibson, D., Alvarez, V., Irving, L., Star, K., Charleer, S., & Verbert, K. (2017). *How to Use Gamified Dashboards and Learning Analytics for Providing Immediate Student Feedback and Performance Tracking in Higher Education*. https://doi.org/10.1145/3041021.3054175

de Hesselle, L. C., Rozgonjuk, D., Sindermann, C., Pontes, H. M., & Montag, C. (2021). *The associations between Big Five personality traits, gaming motives, and self-reported time spent gaming*. In [doi:10.1016/j.paid.2020.110483]. Elsevier Science.

Deci, E. L., & Ryan, R. M. (1980). The Empirical Exploration of Intrinsic Motivational Processes11Preparation of this chapter was facilitated by Research Grant MH 28600 from the National Institute of Mental Health to the first author. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 13, pp. 39-80). Academic Press. https://doi.org/https://doi.org/10.1016/S0065-2601(08)60130-6

Demetriou, C., Özer, B., & Essau, C. (2015). Self-Report Questionnaires. In. https://doi.org/10.1002/9781118625392.wbecp507

Demetrovics, Z., Urbán, R., Nagygyörgy, K., Farkas, J., Zilahy, D., Mervó, B., . . . Harmath, E. (2011). Why do you play? The development of the motives for online gaming questionnaire (MOGQ). *Behavior Research Methods*, *43*(3), 814-825. https://doi.org/10.3758/s13428-011-0091-y

Denisova, A., & Cook, E. (2019). *Power-Ups in Digital Games: The Rewarding Effect of Phantom Game Elements on Player Experience* Proceedings of the Annual Symposium on Computer-Human Interaction in Play, Barcelona, Spain. https://doi.org/10.1145/3311350.3347173

Deterding, S. (2015). The Lens of Intrinsic Skill Atoms: A Method for Gameful Design. *Human-Computer Interaction*, *30*, 294-335. https://doi.org/10.1080/07370024.2014.993471

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). *From Game Design Elements to Gamefulness: Defining Gamification* (Vol. 11). https://doi.org/10.1145/2181037.2181040

Devaux, M., & Sassi, F. (2015). Social disparities in hazardous alcohol use: Self-report bias may lead to incorrect estimates. *European journal of public health*, *26*. https://doi.org/10.1093/eurpub/ckv190

Di Blasi, M., Giardina, A., Giordano, C., Lo Coco, G., Tosto, C., Billieux, J., & Schimmenti, A. (2019). Problematic video game use as an emotional coping strategy: Evidence from a sample of MMORPG gamers. *Journal of Behavioral Addictions*, *8*, 25-34. https://doi.org/10.1556/2006.8.2019.02

Dieris-Hirche, J., Pape, M., te Wildt, B. T., Kehyayan, A., Esch, M., Aicha, S., . . . Bottel, L. (2020). *Problematic gaming behavior and the personality traits of video gamers: A cross-sectional survey* [doi:10.1016/j.chb.2020.106272].

Drachen, A., Canossa, A., & Sorensen, J. (2013). Gameplay Metrics in Game User Research: Examples from the Trenches. In (pp. 285-319). https://doi.org/10.1007/978-1-4471-4769-5_14

Ducheneaut, N., Yee, N., Nickell, E., & Moore, R. (2007). *The life and death of online gaming communities: A look at guilds in World of Warcraft*. https://doi.org/10.1145/1240624.1240750

Eid, M., Riemann, R., Angleitner, A., & Borkenau, P. (2003). Sociability and positive emotionality: Genetic and environmental contributions to the covariation between different facets of extraversion. In (pp. 319-346): Blackwell Publishing.

El-Telbany, O., & Elragal, A. (2017). Gamification of Enterprise Systems: A Lifecycle Approach. *Procedia Computer Science*, *121*, 106-114. https://doi.org/https://doi.org/10.1016/j.procs.2017.11.015

Ensslin, A. (2015). Discourse of Games [https://doi.org/10.1002/9781118611463.wbielsi154]. *The International Encyclopedia of Language and Social Interaction*, 1-6. https://doi.org/https://doi.org/10.1002/9781118611463.wbielsi154 (Major Reference Works)

Esteves, J., Valogianni, K., & Greenhill, A. (2021). Online social games: The effect of social comparison elements on continuance behaviour. *Information & Management*, *58*(4), 103452. https://doi.org/https://doi.org/10.1016/j.im.2021.103452

Everitt, B. S., & Skrondal, A. (2010). The Cambridge dictionary of statistics.

Ewoldsen, D., Eno, C., Okdie, B., Velez, J., Guadagno, R., & DeCoster, J. (2012). Effect of Playing Violent Video Games Cooperatively or Competitively on Subsequent Cooperative Behavior. *CyberPsychology & Behavior*, *15*, 277-280. https://doi.org/10.1089/cyber.2011.0308

Faas, D., Bao, Q., Frey, D., & Yang, M. (2014). The influence of immersion and presence in early stage engineering designing and building. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, *28*, 139-151. https://doi.org/10.1017/S0890060414000055

Falk, A., Gächter, S., & Kovács, J. (1999). Intrinsic motivation and extrinsic incentives in a repeated game with incomplete contracts. *Journal of Economic Psychology*, *20*(3), 251-284. https://doi.org/https://doi.org/10.1016/S0167-4870(99)00009-4

Ferguson, C. J., & Colwell, J. (2017). Understanding Why Scholars Hold Different Views on the Influences of Video Games on Public Health [https://doi.org/10.1111/jcom.12293]. *Journal of Communication*, *67*(3), 305-327. https://doi.org/https://doi.org/10.1111/jcom.12293

Ferguson, E., Zhao, K., & Smillie, L. D. (2020). Personality, preferences and socioeconomic behavior. In *The Cambridge handbook of personality psychology, 2nd ed.* (pp. 477-493): Cambridge University Press.

Ferreira-Oliveira, A., Araújo, A., Fernandes, S. R., & Miguel, I. (2017). *Gamification in the Workplace: A Systematic Literature Review*. https://doi.org/10.1007/978-3-319-56541-5_29

Ferro, L. S. (2018). An analysis of players' personality type and preferences for game elements and mechanics. *Entertainment Computing*, *27*, 73-81. https://doi.org/https://doi.org/10.1016/j.entcom.2018.03.003

Festinger, L. (1954). A Theory of Social Comparison Processes. *Human Relations*, *7*(2), 117-140. https://doi.org/10.1177/001872675400700202

Fitz-Walter, Z., Wyeth, P., Tjondronegoro, D., & Scott-Parker, B. (2013a). *Driven to drive: designing Gamification for a learner logbook smartphone application* Proceedings of the First International Conference on Gameful Design, Research, and Applications, Toronto, Ontario, Canada. https://doi.org/10.1145/2583008.2583014

Fitz-Walter, Z., Wyeth, P., Tjondronegoro, D., & Scott-Parker, B. (2013b). *Driven to drive: designing Gamification for a learner logbook smartphone application* Proceedings of the First International Conference on Gameful Design, Research, and Applications, Toronto, Ontario, Canada. https://doi.org/10.1145/2583008.2583014

Flaounas, I. N., Sudhahar, S., Lansdall-Welfare, T., Hensiger, E., & Cristianini, N. (2013). Big Data Analysis of News and Social Media Content.

Fleming, M. J., & Rick Wood, D. J. (2001). Effects of Violent Versus Nonviolent Video Games on Children's Arousal, Aggressive Mood, and Positive Mood [https://doi.org/10.1111/j.1559-1816.2001.tb00163.x]. *Journal of Applied Social Psychology*, *31*(10), 2047-2071. https://doi.org/https://doi.org/10.1111/j.1559-1816.2001.tb00163.x

Flesch , R. (1979). *How to Write Plain English: A Book for Lawyers and Consumers*. Barnes & Noble.

Fokkema, M., & Greiff, S. *How performing PCA and CFA on the same data equals trouble: Overfitting in the assessment of internal structure and some editorial thoughts on it* Hogrefe Publishing. https://doi.org/10.1027/1015-5759/a000460

Fong, M., Zhao, K., & Smillie, L. (2020). Personality and competitiveness: Extraversion, agreeableness, and their aspects, predict self-reported competitiveness and competitive bidding in experimental auctions. *Personality and Individual Differences*, *169*, 109907. https://doi.org/10.1016/j.paid.2020.109907

Forkosh, J., & Drake, J. (2017). Coloring versus drawing: Effects of cognitive demand on mood repair, flow, and enjoyment [doi:10.1080/07421656.2017.1327272]. *34*,

75-82. https://doi.org/10.1080/07421656.2017.1327272

Franken, R. E., & Brown, D. J. (1995). *Why do people like competition? The motivation for winning, putting forth effort, improving one's performance, performing well, being instrumental, and expressing forceful/aggressive behavior*. In (Version 2) [doi:10.1016/0191-8869(95)00035-5]. Elsevier Science.

Gaetan, S., Bonnet, A., & Pedinielli, J. L. (2012a). Self-perception and life satisfaction in video game addiction in young adolescents (11-14 years old). *L'Encéphale*, *38*, 512-518. https://doi.org/10.1016/j.encep.2012.04.003

Gaetan, S., Bonnet, A., & Pedinielli, J. L. (2012b). Self-perception and life satisfaction in video game addiction in young adolescents (11-14 years old). *L'Encéphale*, *38*, 512-518. https://doi.org/10.1016/j.encep.2012.04.003

Galesic, M., & Bosnjak, M. (2009). Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey. *Public Opinion Quarterly*, *73*. https://doi.org/10.1093/poq/nfp031

Gallo, A., Baumann, P., Agu, E., & Claypool, M. (2016). Exerwalls–an Exercise Alternative to Paywalls in Mobile Games.

Garvey, C. (1990). *Play, Enlarged ed*. Harvard University Press.

Gentile, D. A., Anderson, C. A., Yukawa, S., Ihori, N., Saleem, M., Ming, L. K., . . . Sakamoto, A. (2009). The effects of prosocial video games on prosocial behaviors: International evidence from correlational, longitudinal, and experimental studies. In (pp. 752-763): Sage Publications.

Ghuman, D., & Griffiths, M. D. (2012). A cross-genre study of online gaming: Player demographics, motivation for play, and social interactions among players. *International Journal of Cyber Behavior, Psychology and Learning*, *2*(1), 13-29.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, *48*(1), 26-34. https://doi.org/10.1037/0003-066X.48.1.26

Gourville, J. T., & Soman, D. (2005). Overchoice and Assortment Type: When and Why Variety Backfires. *Marketing Science*, *24*(3), 382-395. https://doi.org/10.1287/mksc.1040.0109

Graham, L., & Gosling, S. (2013). Personality Profiles Associated with Different Motivations for Playing World of Warcraft. *Cyberpsychology, Behavior, and Social Networking*, *16*. https://doi.org/10.1089/cyber.2012.0090

Granic, I., Lobel, A., & Engels, R. C. M. E. (2014). The benefits of playing video games. In

(pp. 66-78): American Psychological Association.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature*, *423*(6939), 534-537. https://doi.org/10.1038/nature01647

Green, C. S., & Bavelier, D. (2007). Action-Video-Game Experience Alters the Spatial Resolution of Vision. *Psychological Science*, *18*(1), 88-94. https://doi.org/10.1111/j.1467-9280.2007.01853.x

Greenberg, B. S., Sherry, J., Lachlan, K., Lucas, K., & Holmstrom, A. (2010). *Orientations to video games among gender and age groups*. In (Version 2) [doi:10.1177/1046878108319930]. Sage Publications.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95-112. https://doi.org/10.1007/BF02289823

Greitemeyer, T., & Osswald, S. (2009). Prosocial video games reduce aggressive cognitions. *Journal of Experimental Social Psychology*, *45*(4), 896-900. https://doi.org/10.1016/j.jesp.2009.04.005

Greitemeyer, T., Osswald, S., & Brauer, M. (2010). Playing prosocial video games increases empathy and decreases schadenfreude. In (pp. 796-802): American Psychological Association.

Griffiths, M., & Nuyens, F. (2017). An Overview of Structural Characteristics in Problematic Video Game Playing. *Current Addiction Reviews*, *4*. https://doi.org/10.1007/s40429-017-0162-y

Gylfason, H. F., Halldorsson, F., & Kristinsson, K. (2016). Personality in Gneezy's cheap talk game: The interaction between Honesty-Humility and Extraversion in predicting deceptive behavior. *Personality and Individual Differences*, *96*, 222-226. https://doi.org/10.1016/j.paid.2016.02.075

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2014). Pearson new international edition. *Multivariate data analysis, Seventh Edition. Pearson Education Limited Harlow, Essex*.

Hakulinen, L., Auvinen, T., & Korhonen, A. (2015). The Effect of Achievement Badges on Students' Behavior: An Empirical Study in a University-Level Computer Science Course. *International Journal of Emerging Technologies in Learning (iJET)*, *10*. https://doi.org/10.3991/ijet.v10i1.4221

Hamari, J., & Koivisto, J. (2014). Measuring flow in Gamification: Dispositional Flow Scale-2. *Computers in Human Behavior*, *40*, 133-143.

https://doi.org/10.1016/j.chb.2014.07.048

Hamari, J., Koivisto, J., & Sarsa, H. (2014). *Does Gamification Work? — A Literature Review of Empirical Studies on Gamification*. https://doi.org/10.1109/HICSS.2014.377

Hamari, J., & Lehdonvirta, V. (2010). Game design as marketing: How game mechanics create demand for virtual goods. *International Journal of Business Science and Applied Management*, *5*.

Hamari, J., & Tuunanen, J. (2014). Player Types: A Meta-synthesis. *Transactions of the Digital Games Research Association*, *1*, 29-53. https://doi.org/10.26503/todigra.v1i2.13

Harris, M. A., Brett, C. E., Johnson, W., & Deary, I. J. (2016). Personality stability from age 14 to age 77 years. *Psychology and Aging*, *31*(8), 862-874. https://doi.org/10.1037/pag0000133

Hartup, P. (2015). *It's OK for PC gamers to be a little arrogant – they know they've backed the right platform*. TheNewStatesmen. https://www.newstatesman.com/culture/2015/01/it-s-ok-pc-gamers-be-little-arrogant-they-know-theyve-backed-right-platform

Haworth, R., Bostani, S., & Sedig, K. (2010). Visualizing Decision Trees in Games to Support Children's Analytic Reasoning: Any Negative Effects on Gameplay? *International Journal of Computer Games Technology*, *2010*. https://doi.org/10.1155/2010/578784

Heifetz, A., Shannon, C., & Spiegel, Y. (2007). The Dynamic Evolution of Preferences. *Economic Theory*, *32*(2), 251-286. https://doi.org/10.1007/s00199-006-0121-7

Heller, G. M. (2012). *Psychopathology, adaptive functioning, and motivational needs as predictors of video game use and engagement* ProQuest Information & Learning]. psyh. US. http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2012-99140-560

Hennighausen, C., Hudders, L., Lange, B. P., & Fink, H. (2016). *What if the rival drives a porsche?: Luxury car spending as a costly signal in male intrasexual competition*. In (Version 4) [doi:10.1177/1474704916678217]. Sage Publications.

Holden, C. J., Zeigler-Hill, V., Pham, M. N., & Shackelford, T. K. (2014). Personality features and mate retention strategies: Honesty–humility and the willingness to manipulate, deceive, and exploit romantic partners. *Personality and Individual*

Differences, *57*, 31-36. https://doi.org/10.1016/j.paid.2013.09.018

Hsieh, C.-C., Lin, P.-S., Hsu, W.-C., Wang, J.-S., Huang, Y.-C., Lim, A.-Y., & Hsu, Y.-C. (2018). The Effectiveness of a Virtual Reality-Based Tai Chi Exercise on Cognitive and Physical Function in Older Adults with Cognitive Impairment. *Dementia and Geriatric Cognitive Disorders*, *46*, 358-370. https://doi.org/10.1159/000494659

Hsu, M. (2017). Neuromarketing: Inside the Mind of the Consumer. *California Management Review*, *59*(4), 5-22. https://doi.org/10.1177/0008125617720208

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Huber, G. P. (1981). The Nature of Organizational Decision Making and the Design of Decision Support Systems. *MIS Quarterly*, *5*(2), 1-10. https://doi.org/10.2307/249220

Hudders, L., De Backer, C., Fisher, M., & Vyncke, P. (2014). *The rival wears Prada: Luxury consumption as a female competition strategy*. In (Version 3) [doi:10.1177/14747049140120030 6]. Idium AS.

Hui, S. L. T., & See, S. L. (2015). Enhancing User Experience Through Customisation of UI Design. *Procedia Manufacturing*, *3*, 1932-1937. https://doi.org/https://doi.org/10.1016/j.promfg.2015.07.237

Hunter, D., & Werbach, K. (2012). *For the win* (Vol. 2). Wharton digital press Philadelphia, PA, USA.

Hunter, L. (2012). Challenging the reported disadvantages of e-questionnaires and addressing methodological issues of online data collection. *Nurse researcher*, *20*, 11-20. https://doi.org/10.7748/nr2012.09.20.1.11.c9303

Huotari, K., & Hamari, J. (2012). *Defining Gamification - A Service Marketing Perspective*. https://doi.org/10.1145/2393132.2393137

Hussain, U., Jabarkhail, S., Cunningham, G. B., & Madsen, J. A. (2021). The dual nature of escapism in video gaming: A meta-analytic approach. *Computers in Human Behavior Reports*, *3*, 100081. https://doi.org/https://doi.org/10.1016/j.chbr.2021.100081

Högberg, J., Hamari, J., & Wästlund, E. (2019). Gameful Experience Questionnaire (GAMEFULQUEST): an instrument for measuring the perceived gamefulness of system use. *User Modeling and User-Adapted Interaction*, *29*.

https://doi.org/10.1007/s11257-019-09223-w

Jang, H., Kim, E. J., & Reeve, J. (2016). Why students become more engaged or more disengaged during the semester: A self-determination theory dual-process model. *Learning and Instruction*, *43*, 27-38. https://doi.org/https://doi.org/10.1016/j.learninstruc.2016.01.002

Jeong, H., Yim, H., Lee, S.-Y., Lee, H., Potenza, M., Kwon, J.-H., . . . Choi, J.-S. (2018). Discordance between self-report and clinical diagnosis of Internet gaming disorder in adolescents. *Scientific Reports*, *8*. https://doi.org/10.1038/s41598-018-28478-8

Jia, Y., Xu, B., Masterson, Y., & Voida, S. (2016). *Personality-targeted Gamification: A Survey Study on Personality Traits and Motivational Affordances*. https://doi.org/10.1145/2858036.2858515

Johannes, N., Vuorre, M., & Przybylski, A. (2020). *Video game play is positively correlated with well-being*. https://doi.org/10.31234/osf.io/qrjza

Johnson, D., Deterding, S., Kuhn, K.-A., Staneva, A., Stoyanov, S., & Hides, L. (2016). Gamification for health and wellbeing: A systematic review of the literature. *Internet Interventions*, *6*, 89-106. https://doi.org/https://doi.org/10.1016/j.invent.2016.10.002

Jones, B. A., Madden, G. J., Wengreen, H. J., Aguilar, S. S., & Desjardins, E. A. (2014). Gamification of Dietary Decision-Making in an Elementary-School Cafeteria. *PLOS ONE*, *9*(4), e93872. https://doi.org/10.1371/journal.pone.0093872

Juul, J. (2011). *Half-real: Video games between real rules and fictional worlds*. MIT press.

Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis [doi:10.1007/BF02289343]. *34*, 183-202. https://doi.org/10.1007/BF02289343

Kachelmeier, S. J., & Towry, K. L. (2005). The Limitations of Experimental Design: A Case Study Involving Monetary Incentive Effects in Laboratory Markets. *Experimental Economics*, *8*(1), 21-33. https://doi.org/10.1007/s10683-005-0435-5

Kahn, A. S., Shen, C., Lu, L., Ratan, R. A., Coary, S., Hou, J., . . . Williams, D. (2015). The Trojan Player Typology: A cross-genre, cross-cultural, behaviorally validated scale of video game play motivations. *Computers in Human Behavior*, *49*, 354-361. https://doi.org/10.1016/j.chb.2015.03.018

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31-36.

https://doi.org/10.1007/BF02291575

Kardefelt-Winther, D. (2014). A conceptual and methodological critique of internet addiction research: Towards a model of compensatory internet use. *Computers in Human Behavior*, *31*, 351-354. https://doi.org/https://doi.org/10.1016/j.chb.2013.10.059

Kaushal, R., & Kwantes, C. T. (2006). The role of culture and personality in choice of conflict management strategy. *International Journal of Intercultural Relations*, *30*(5), 579-603. https://doi.org/https://doi.org/10.1016/j.ijintrel.2006.01.001

Kaye, L., Kowert, R., & Quinn, S. (2018). *The role of social identity and online social capital on psychosocial outcomes in MMO players*. https://doi.org/10.31219/osf.io/32uvr

Kelloway, E. K. (1995). Structural equation modelling in perspective [https://doi.org/10.1002/job.4030160304]. *Journal of Organizational Behavior*, *16*(3), 215-224. https://doi.org/https://doi.org/10.1002/job.4030160304

Kent, S. L. (2010). *The Ultimate History of Video Games, Volume 1: From Pong to Pokemon and Beyond... the Story Behind the Craze That Touched Our Lives and Changed the World* (Vol. 1). Crown.

Kiili, K., Moeller, K., & Ninaus, M. (2018). Evaluating the effectiveness of a game-based rational number training - In-game metrics as learning indicators. *Computers & Education*, *120*, 13-28. https://doi.org/https://doi.org/10.1016/j.compedu.2018.01.012

Kim, J., Lee, A., & Ryu, H. (2013). Personality and its effects on learning performance: Design guidelines for an adaptive e-learning system based on a user model. *International Journal of Industrial Ergonomics*, *43*, 450–461. https://doi.org/10.1016/j.ergon.2013.03.001

Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel*.

King, D. L., Delfabbro, P. H., Perales, J. C., Deleuze, J., Király, O., Krossbakken, E., & Billieux, J. (2019). Maladaptive player-game relationships in problematic gaming and gaming disorder: A systematic review. *Clinical Psychology Review*, *73*, 101777. https://doi.org/https://doi.org/10.1016/j.cpr.2019.101777

Klock, A., Gasparini, I., Pimenta, M., & Palazzo Moreira de Oliveira, J. (2015). "Everybody is playing the game, but nobody's rules are the same": Towards adaptation of

Gamification based on users' characteristics. *Bulletin of the Technical Committee on Learning Technology*, *17*, 22.

Knee, C. R., & Neighbors, C. (2002). *Self-determination, perception of peer pressure, and drinking among college students* [doi:10.1111/j.1559-1816.2002.tb00228.x].

Kofinas, A. K., & Tsay, C. H.-H. (2021). In favor of large classes: a social networks perspective on experiential learning. *Journal of Management Education*, *45*(5), 760-785.

Koivisto, J., & Hamari, J. (2019). The rise of the motivational information systems: A review of Gamification research. *International Journal of Information Management*, *45*, 210. https://doi.org/10.1016/j.ijinfomgt.2018.10.013

Kolo, C., & Baur, T. (2004). Living a virtual life: Social dynamics of online gaming. *Game studies*, *4*(1), 1-31.

Kuss, D. J., van Rooij, A. J., Shorter, G. W., Griffiths, M. D., & van de Mheen, D. (2013). Internet addiction in adolescents: Prevalence and risk factors. In

Lafrenière, M.-A., Verner-Filion, J., & Vallerand, R. (2012). Development and validation of the Gaming Motivation Scale (GAMS). *Personality and Individual Differences*, *53*, 827–831. https://doi.org/10.1016/j.paid.2012.06.013

Lahti, D. C. (2015). The Limits of Artificial Stimuli in Behavioral Research: The Umwelt Gamble [https://doi.org/10.1111/eth.12361]. *Ethology*, *121*(6), 529-537. https://doi.org/https://doi.org/10.1111/eth.12361

Lameras, P., Arnab, S., Dunwell, I., Stewart, C., Clarke, S., & Petridis, P. (2017). Essential features of serious games design in higher education: Linking learning attributes to game mechanics [https://doi.org/10.1111/bjet.12467]. *British Journal of Educational Technology*, *48*(4), 972-994. https://doi.org/https://doi.org/10.1111/bjet.12467

Landers, R. N., & Landers, A. K. (2014). An Empirical Test of the Theory of Gamified Learning. *Simulation & Gaming*, *45*(6), 769-785. https://doi.org/10.1177/1046878114563662

Le Hénaff, B., Michinov, N., Le Bohec, O., & Delaval, M. (2015). Social gaming is inSIDE: Impact of anonymity and group identity on performance in a team game-based learning environment. In (pp. 84-95): Elsevier Science.

Lee, J., Lee, M., & Choi, I. (2012). Social Network Games Uncovered: Motivations and Their Attitudinal and Behavioral Outcomes. *Cyberpsychology, behavior and social*

networking, *15*. https://doi.org/10.1089/cyber.2012.0093

Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO personality inventory. *Multivariate Behavioral Research*, *39*(2), 329-358. https://doi.org/10.1207/s15327906mbr3902_8

Lee, K., Ashton, M. C., & de Vries, R. E. (2005). Predicting Workplace Delinquency and Integrity with the HEXACO and Five-Factor Models of Personality Structure. *Human Performance*, *18*(2), 179-197. https://doi.org/10.1207/s15327043hup1802_4

Lehenbauer-Baum, M., Klaps, A., Kovacovsky, Z., Witzmann, K., Zahlbruckner, R., & Stetina, B. (2015). Addiction and Engagement: An Explorative Study Toward Classification Criteria for Internet Gaming Disorder. *Cyberpsychology, behavior and social networking*, *18*, 343-349. https://doi.org/10.1089/cyber.2015.0063

Leigh, J. S. A., Steckler, E., & Ennis, W. (2018). Harnessing the Potential of Gamification in Humanistic Management Teaching: Games for Change. *Humanistic Management: Social Entrepreneurship and Mindfulness, Volume II: Foundations, Cases, and Exercises*.

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology*, *22*(6), 1027-1038. https://doi.org/https://doi.org/10.1016/j.conb.2012.06.001

Lichtenberk, F. (2002). The Possessive-Benefactive Connection. *Oceanic Linguistics*, *41*(2), 439-474. https://doi.org/10.2307/3623316

Liu, C., Agrawal, P., Sarkar, N., & Chen, S. (2009). Dynamic Difficulty Adjustment in Computer Games Through Real-Time Anxiety-Based Affective Feedback. *International Journal of Human–Computer Interaction*, *25*(6), 506-529. https://doi.org/10.1080/10447310902963944

Lopez, C. E., & Tucker, C. S. (2019). The effects of player type on performance: A Gamification case study. *Computers in Human Behavior*, *91*, 333-345. https://doi.org/https://doi.org/10.1016/j.chb.2018.10.005

Lounis, S., Neratzouli, X., & Pramatari, K. (2013). *Can Gamification Increase Consumer Engagement? A Qualitative Approach on a Green Case* (Vol. 399). https://doi.org/10.1007/978-3-642-37437-1_17

Lucas, R. E. (2018). Reevaluating the strengths and weaknesses of self-report measures of subjective well-being. *Handbook of well-being. Salt Lake City, UT: DEF Publishers*.

Lyubomirsky, S., & Ross, L. (1997). Hedonic consequences of social comparison: A contrast of happy and unhappy people. In (pp. 1141-1157): American Psychological Association.

Macey, J., & Hamari, J. (2018). eSports, skins and loot boxes: Participants, practices and problematic behaviour associated with emergent forms of gambling. *New Media & Society*, *21*(1), 20-41. https://doi.org/10.1177/1461444818786216

Major, D. A., Turner, J. E., & Fletcher, T. D. (2006). Linking proactive personality and the Big Five to motivation to learn and development activity. *Journal of Applied Psychology*, *91*(4), 927-935. https://doi.org/10.1037/0021-9010.91.4.927

Marczewski, A. (2015). User Types HEXAD. In (pp. 65-80).

Marino, C., Canale, N., Vieno, A., Caselli, G., Scacchi, L., & Spada, M. M. (2020). Social anxiety and Internet gaming disorder: The role of motives and metacognitions. *Journal of Behavioral Addictions*, *9*(3), 617-628. https://doi.org/10.1556/2006.2020.00044

Martí-Parreño, J., Bermejo-Berros, J., & Aldás-Manzano, J. (2017). Product Placement in Video Games: The Effect of Brand Familiarity and Repetition on Consumers' Memory. *Journal of Interactive Marketing*, *38*(C), 55-63.

Matthews, G., & Gilliland, K. (1999). The personality theories of H.J. Eysenck and J.A. Gray: a comparative review. *Personality and Individual Differences*, *26*(4), 583-626. https://doi.org/https://doi.org/10.1016/S0191-8869(98)00158-5

Mazlan, M. N. A., & Bakar, M. A. A. (2013). Students' Perception of Self-presentation towards Avatar. *Procedia - Social and Behavioral Sciences*, *97*, 361-367. https://doi.org/https://doi.org/10.1016/j.sbspro.2013.10.246

McCarthy, C. (2019). Address students' self-perception of online invincibility, anonymity [https://doi.org/10.1002/say.30623]. *Student Affairs Today*, *22*(4), 1-4. https://doi.org/https://doi.org/10.1002/say.30623

McGonigal, J. (2011). Reality is broken: Why games make us better and how they can change the world. In *Reality is broken: Why games make us better and how they can change the world.* (pp. 388-388): Penguin Press.

McMahan, C. S., Wang, L., & Tebbs, J. M. (2013). Regression analysis for current status data using the EM algorithm [https://doi.org/10.1002/sim.5863]. *Statistics in Medicine*, *32*(25), 4452-4466. https://doi.org/https://doi.org/10.1002/sim.5863

Mengel, F. (2014). Computer Games and Prosocial Behaviour. *PLOS ONE*, *9*(4), e94099. https://doi.org/10.1371/journal.pone.0094099

Merikivi, J., Tuunainen, V., & Nguyen, D. (2017). What makes continued mobile gaming enjoyable? *Computers in Human Behavior*, *68*, 411-421. https://doi.org/https://doi.org/10.1016/j.chb.2016.11.070

Michailidis, L., Balaguer-Ballester, E., & He, X. (2018). Flow and Immersion in Video Games: The Aftermath of a Conceptual Challenge. *Frontiers in Psychology*, *9*, 1682. https://doi.org/10.3389/fpsyg.2018.01682

Misuraca, R., Teuscher, U., & Faraci, P. (2016). Is more choice always worse? Age differences in the overchoice effect. *Journal of Cognitive Psychology*, *28*(2), 242-255. https://doi.org/10.1080/20445911.2015.1118107

Moore, C., Barbour, K., & Lee, K. (2017). Five Dimensions of Online Persona. *Persona Studies*, *3*, 1-12. https://doi.org/10.21153/ps2017vol3no1art658

Morling, B., & Lee, J. M. (2017). Culture and motivation. In *The Praeger handbook of personality across cultures: Culture and characteristic adaptations, Vol. 2* (pp. 61-89): Praeger/ABC-CLIO.

Mueller, G., & Plug, E. (2006). Estimating the Effect of Personality on Male and Female Earnings. *ILR Review*, *60*(1), 3-22. https://doi.org/10.1177/001979390606000101

Muntean, C. I. (2011). Raising engagement in e-learning through Gamification.

Müller, K., Beutel, M., Egloff, B., & Wölfling, K. (2014). Investigating Risk Factors for Internet Gaming Disorder: A Comparison of Patients with Addictive Gaming, Pathological Gamblers and Healthy Controls regarding the Big Five Personality Traits. *European addiction research*, *20*, 129-136. https://doi.org/10.1159/000355832

Nacke, L. E. (2018). Introduction to biometric measures for Games User Research. In (pp. 281-299). https://doi.org/10.1093/oso/9780198794844.003.0016

Nacke, L. E., & Deterding, S. (2017). *The maturing of Gamification research*. In [doi:10.1016/j.chb.2016.11.062]. Elsevier Science.

Napierala, M. A. (2012). What is the Bonferroni correction?. Aaos Now, 40-41.

Nyhus, E., & Pons, E. (2012). Personality and the Gender Wage Gap. *Applied Economics*, *44*, 105-118. https://doi.org/10.1080/00036846.2010.500272

Oinas-Kukkonen, H., & Harjumaa, M. (2009). Persuasive Systems Design: Key Issues, Process Model, and System Features. *Communications of the Association for Information Systems*, *24*. https://doi.org/10.17705/1CAIS.02428

Omar, A., Mohamad Ali, N., & Judi, H. (2011). *Measuring flow experience in gaming devices interactions*. https://doi.org/10.1109/ICEEI.2011.6021636

Orji, R., Tondello, G., & Nacke, L. (2018). *Personalizing Persuasive Strategies in Gameful Systems to Gamification User Types*. https://doi.org/10.1145/3173574.3174009

Orji, R., Vassileva, J., & Mandryk, R. L. (2014). Modeling the efficacy of persuasive strategies for different gamer types in serious games for health. *User Modeling and User-Adapted Interaction*, *24*(5), 453-498. https://doi.org/10.1007/s11257-014-9149-8

O'Brien, H. L. (2010). The influence of hedonic and utilitarian motivations on user engagement: The case of online shopping experiences. *Interacting with Computers*, *22*(5), 344-352. https://doi.org/https://doi.org/10.1016/j.intcom.2010.04.001

P.S, G. (2015). Platform Comparison Between Games Console, Mobile Games And PC Games. *SISFORMA*, *2*, 23. https://doi.org/10.24167/sisforma.v2i1.407

Paik, P., & Shi, C.-K. (2013). Playful gender swapping: User attitudes toward gender in MMORPG avatar customisation. *Digital Creativity*, *24*. https://doi.org/10.1080/14626268.2013.767275

Peever, N., Johnson, D., & Gardner, J. (2012). *Personality & video game genre preferences* Proceedings of The 8th Australasian Conference on Interactive Entertainment: Playing the System, Auckland, New Zealand. https://doi.org/10.1145/2336727.2336747

Petry, N. M., Rehbein, F., Ko, C.-H., & O'Brien, C. P. (2015). Internet Gaming Disorder in the DSM-5. *Current Psychiatry Reports*, *17*(9), 72. https://doi.org/10.1007/s11920-015-0610-0

Pluymen, P., Pequet, A., Thomas, H., & Warnell, K. R. (2021). Capturing individual differences in social motivation using a novel interactive task. *Personality and Individual Differences*, *177*, 110725. https://doi.org/https://doi.org/10.1016/j.paid.2021.110725

Prensky, M. (2012). *From digital natives to digital wisdom: Hopeful essays for 21st century learning*. https://doi.org/10.4135/9781483387765

Preston, J. (2012). Absorbed in video game play: From immersion to the development of consciousness. *Video game play and consciousness.*, 143-170. (Perspectives on cognitive psychology.)

Pringle, H. M. (2015). Conjuring the ideal self: An investigation of self-presentation in

video game avatars. *Press Start*, *2*(1), 1-20.

Przybylski, A. K., Rigby, C. S., & Ryan, R. M. (2010). *A motivational model of video game engagement*. In (Version 2) [doi:10.1037/a0019440]. Educational Publishing Foundation.

Raykov, T., & Marcoulides, G. A. (1999). On desirability of parsimony in structural equation model selection. In (Vol. 6, pp. 292-300). US: Lawrence Erlbaum.

Rehbein, F., Staudt, A., Hanslmaier, M., & Kliem, S. (2015). Video game playing in the general adult population of Germany: Can higher gaming time of males be explained by gender specific genre preferences? *Computers in Human Behavior*, 729-735. https://doi.org/10.1016/j.chb.2015.10.016

Reynolds, L., Schwanda Sosik, V., & Cosley, D. (2013). When Wii doesn't Fit: how non-beginners react to Wii Fit's Gamification. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/2583008.2583027

Rigby, S. (2004). Player Motivational Analysis: A model for applied research into the motivational dynamics of virtual worlds. *Motivation Research Group, University of Rochester, Rochester, NY*.

Rocha, J. B., Mascarenhas, S. F., & Prada, R. (2008). Game Mechanics for Cooperative Games.

Rodriguez-Guerrero, C., Knaepen, K., Fraile-Marinero, J. C., Pérez-Turiel, J., Gonzalez-de-Garibay, V., & Lefeber, D. (2017). Improving Challenge/Skill Ratio in a Multimodal Interface by Simultaneously Adapting Game Difficulty and Haptic Assistance through Psychophysiological and Performance Feedback. *Frontiers in Neuroscience*, *11*.

Rogers, E. M., & Larsen, J. K. (1984). *Silicon Valley fever: Growth of high-technology culture* (Vol. 112). Basic books New York.

Ross, S. R., Canada, K. E., & Rausch, M. K. (2002). Self-handicapping and the Five Factor Model of personality: Mediation between neuroticism and conscientiousness. *Personality and Individual Differences*, *32*(7), 1173-1184. https://doi.org/10.1016/S0191-8869(01)00079-4

Rosser, J. C., Jr., Lynch, P. J., Cuddihy, L., Gentile, D. A., Klonsky, J., & Merrell, R. (2007). The Impact of Video Games on Training Surgeons in the 21st Century. *Archives of Surgery*, *142*(2), 181-186. https://doi.org/10.1001/archsurg.142.2.181

Rubin, V., & Camm, S. (2013). Deception in video games: Examining varieties of griefing.

Online Information Review, *37*, 369-387. https://doi.org/10.1108/OIR-10-2011-0181

Sackett, P. R., & Walmsley, P. T. (2014). Which personality attributes are most important in the workplace? *Perspectives on Psychological Science*, *9*(5), 538-551. https://doi.org/10.1177/1745691614543972

Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017a). *How Gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction*. In [doi:10.1016/j.chb.2016.12.033]. Elsevier Science.

Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017b). How Gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior*, *69*, 371-380. https://doi.org/https://doi.org/10.1016/j.chb.2016.12.033

Sailer, M., & Homner, L. (2020). The Gamification of Learning: a Meta-analysis. *Educational Psychology Review*, *32*, 77-112. https://doi.org/10.1007/s10648-019-09498-w

Salen, K., & Zimmerman, E. (2004). Rules of Play : Game Design Fundamentals / K. Salen, E. Zimmerman ; pról. de Frank Lantz.

Samuels, P. (2017). *Advice on Exploratory Factor Analysis*.

san martín, S., Jimenez Torres, N. H., Camarero, C., & San Jose, R. (2020). The Path between Personality, Self-Efficacy, and Shopping Regarding Games Apps. *Journal of theoretical and applied electronic commerce research*, *15*. https://doi.org/10.4067/S0718-18762020000200105

Sassenrath, C. (2019). "Let me show you how nice I am": Impression management as bias in empathic responses. *Social Psychological and Personality Science*. https://doi.org/10.1177/1948550619884566]

Scott, J., & Porter-Armstrong, A. (2013). Impact of Multiplayer Online Role-Playing Games upon the Psychosocial Well-Being of Adolescents and Young Adults: Reviewing the Evidence. *Psychiatry journal*, *2013*, 464685. https://doi.org/10.1155/2013/464685

Seah, M.-l. (2008). *From Immersion to Addiction in Videogames*. https://doi.org/10.14236/ewic/HCI2008.6

Sedig, K., Parsons, P., & Haworth, R. (2017). Player-Game Interaction and Cognitive Gameplay: A Taxonomic Framework for the Core Mechanic of Videogames. *Informatics*, *4*, 4.

Shenk, D. (2011). *The Immortal Game: A History of Chess; Or How 32 Carved Pieces on a Board Illuminated Our Understanding of War, Art, Science, and the Human Brain*. Anchor Canada.

Shepard, L., Penuel, W., & Pellegrino, J. (2018). Using Learning and Motivation Theories to Coherently Link Formative Assessment, Grading Practices, and Large-Scale Assessment. *Educational Measurement: Issues and Practice*, *37*, 21-34. https://doi.org/10.1111/emip.12189

Sherry, J. L., Lucas, K., Greenberg, B. S., & Lachlan, K. (2006). Video game uses and gratifications as predictors of use and game preference. *Playing video games: Motives, responses, and consequences*, *24*, 213-224.

Shi, Z., Zhou, A., Liu, P., Zhang, P., & Han, W. (2011). An EEG study on the effect of self-relevant possessive pronoun: Self-referential content and first-person perspective. *Neuroscience letters*, *494*, 174-179. https://doi.org/10.1016/j.neulet.2011.03.007

Shibuya, A., Teramoto, M., Shoun, A., & Akiyama, K. (2019). Long-Term Effects of In-Game Purchases and Event Game Mechanics on Young Mobile Social Game Players in Japan. *Simulation & Gaming*, *50*, 104687811881967. https://doi.org/10.1177/1046878118819677

Sidéra, I., & Vornicu, A. (2016). The archaeology of games. Playing with knucklebones in the Early Chalcolithic of the Balkans. In: Verlag Dr. Rudolf Habelt GmbH.

Sigailov-Lanfranchi, J. (2019). Impact of Different Levels of Difficulty on Immersion in Video Games.

Silvia, P. J., & Christensen, A. P. (2020). Looking up at the curious personality: Individual differences in curiosity and openness to experience. *Current Opinion in Behavioral Sciences*, *35*, 1-6. https://doi.org/10.1016/j.cobeha.2020.05.013

Smiderle, R., Rigo, S. J., Marques, L. B., Peçanha de Miranda Coelho, J. A., & Jaques, P. A. (2020). The impact of Gamification on students' learning, engagement and behavior based on their personality traits. *Smart Learning Environments*, *7*(1), 3. https://doi.org/10.1186/s40561-019-0098-x

Smith, B., & Blumstein, D. (2008). Fitness consequences of personality: A meta-analysis. *Behavioral Ecology*, *19*, 448-455. https://doi.org/10.1093/beheco/arm144

Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural equation modeling. In (Vol. 42, pp. 893-898). Netherlands: Elsevier Science.

Stern, W. (1900). *Über Psychologie der individuellen Differenzen: Ideen zu einer" differentiellen Psychologie"*. JA Barth.

Studios, S. (2022). Mobile Game Development Cost: What Are the Factors that Influence the Price? In.

Suh, A., Cheung, C. M. K., Ahuja, M., & Wagner, C. (2017). Gamification in the Workplace: The Central Role of the Aesthetic Experience. *Journal of Management Information Systems*, *34*(1), 268-305. https://doi.org/10.1080/07421222.2017.1297642

Suh, A., Wagner, C., & Liu, L. (2016). Enhancing User Engagement through Gamification. *Journal of Computer Information Systems*, *58*, 1-10. https://doi.org/10.1080/08874417.2016.1229143

Suhr, D. (2006). *Exploratory or Confirmatory Factor Analysis*.

Suler, J. (2004). The Online Disinhibition Effect. *CyberPsychology & Behavior*, *7*(3), 321-326. https://doi.org/10.1089/1094931041291295

Tabachnick, B. G., & Fidell, L. S. (2014). Using Multivariate Statistics: Pearson New International Edition, edn.

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics* (Vol. 5). pearson Boston, MA.

Tekofsky, S., Miller, P., Spronck, P., & Slavin, K. (2017). *The Effect of Gender, Native English Speaking, and Age on Game Genre Preference and Gaming Motivations* (Vol. 178). https://doi.org/10.1007/978-3-319-49616-0_17

Thorne, H., Smith, J., Morgan, P., Babic, M., & Lubans, D. (2014). Video game genre preference, physical activity and screen-time in adolescent boys from low-income communities. *Journal of Adolescence*, *37*, 1345–1352. https://doi.org/10.1016/j.adolescence.2014.09.012

Thurstone, L. L. (1947). Multiple-factor analysis; a development and expansion of The Vectors of Mind.

Tlili, A., Denden, M., Essalmi, F., Jemni, M., Kinshuk, D., Chen, N.-S., & Huang, R. (2019). Does Providing a Personalized Educational Game Based on Personality Matter? A Case Study. *IEEE Access*, *PP*, 1-1. https://doi.org/10.1109/ACCESS.2019.2936384

Tomić, N. (2017). Effects of micro transactions on video games industry. *Megatrend revija*, *14*, 239-257. https://doi.org/10.5937/MegRev1703239T

Tondello, G., Mora, A., & Nacke, L. (2017). *Elements of Gameful Design Emerging from User Preferences*. https://doi.org/10.1145/3116595.3116627

Torres-Marín, J., Proyer, R. T., López-Benítez, R., Brauer, K., & Carretero-Dios, H. (2019). Beyond the big five as predictors of dispositions toward ridicule and being laughed at: The HEXACO model and the dark triad [https://doi.org/10.1111/sjop.12563]. *Scandinavian Journal of Psychology*, *60*(5), 473-483. https://doi.org/https://doi.org/10.1111/sjop.12563

VanderWeele, T. J., & Mathur, M. B. (2019). Some desirable properties of the Bonferroni correction: is the Bonferroni correction really so bad?. American journal of epidemiology, 188(3), 617-618.

van Ommen, M. (2016). Guild Wars 2, the Frankfurt School and Dialectical Fairy Scenes: A Critical Approach Towards Massively Multiplayer Online Video Games. *Games and Culture*, *13*(6), 547-567. https://doi.org/10.1177/1555412015627392

van Roy, R., & Zaman, B. (2017). Why Gamification Fails in Education and How to Make It Successful: Introducing Nine Gamification Heuristics Based on Self-Determination Theory. In (pp. 485-509). https://doi.org/10.1007/978-3-319-51645-5_22

Veloski, J., Tai, S., Evans, A. S., & Nash, D. B. (2005). Clinical Vignette-Based Surveys: A Tool for Assessing Physician Practice Variation. *American Journal of Medical Quality*, *20*(3), 151-157. https://doi.org/10.1177/1062860605274520

Vesa, M., Hamari, J., Harviainen, J. T., & Warmelink, H. (2016). Computer Games and Organization Studies. *Organization Studies*, *38*(2), 273-284. https://doi.org/10.1177/0170840616663242

Walther, B., Morgenstern, M., & Hanewinkel, R. (2012). Co-Occurrence of Addictive Behaviours: Personality Factors Related to Substance Use, Gambling and Computer Gaming. *European Addiction Research*, *18*(4), 167-174. https://doi.org/10.1159/000335662

Wan, C.-S., & Chiou, W.-B. (2006). Psychological Motives and Online Games Addiction: ATest of Flow Theory and Humanistic Needs Theory for Taiwanese Adolescents. *CyberPsychology & Behavior*, *9*(3), 317-324. https://doi.org/10.1089/cpb.2006.9.317

Wan, C.-S., & Chiou, W.-B. (2007). The motivations of adolescents who are addicted to

online games: A cognitive perspective. *Adolescence*, *42*, 179-197.

Wang, C.-W., Ho, R. T. H., Chan, C. L. W., & Tse, S. (2015). Exploring personality characteristics of Chinese adolescents with internet-related addictive behaviors: Trait differences for gaming addiction and social networking addiction. In (pp. 32-35): Elsevier Science.

Wang, C. K. J., Liu, W. C., & Khoo, A. (2009). The Psychometric Properties of Dispositional Flow Scale-2 in Internet Gaming. *Current Psychology*, *28*(3), 194-201. https://doi.org/10.1007/s12144-009-9058-x

Weisberg, Y., DeYoung, C., & Hirsh, J. (2011). Gender Differences in Personality across the Ten Aspects of the Big Five [Original Research]. *Frontiers in Psychology*, *2*.

Weller, J. A., & Thulin, E. W. (2012). Do honest people take fewer risks? Personality correlates of risk-taking to achieve gains and avoid losses in HEXACO space. *Personality and Individual Differences*, *53*(7), 923-926. https://doi.org/10.1016/j.paid.2012.06.010

White, B. A. (2014). Who cares when nobody is watching? Psychopathic traits and empathy in prosocial behaviors. *Personality and Individual Differences*, *56*, 116-121. https://doi.org/https://doi.org/10.1016/j.paid.2013.08.033

Willmer, M., Westerberg Jacobson, J., & Lindberg, M. (2019). Exploratory and confirmatory factor analysis of the 9-item Utrecht Work Engagement Scale in a multi-occupational female sample: A cross-sectional study. Switzerland.

Willmer, M., Westerberg Jacobson, J., & Lindberg, M. (2019). Exploratory and confirmatory factor analysis of the 9-item Utrecht Work Engagement Scale in a multi-occupational female sample: A cross-sectional study. Switzerland.

Wilpert, B. (2001). Organizational Behavior, Psychology of. In N. J. Smelser & P. B. Baltes (Eds.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 10917-10921). Pergamon. https://doi.org/https://doi.org/10.1016/B0-08-043076-7/01403-0

Witkowski, J., & Kiba-Janiak, M. (2014). The Role of Local Governments in the Development of City Logistics. *Procedia - Social and Behavioral Sciences*, *125*, 373-385. https://doi.org/https://doi.org/10.1016/j.sbspro.2014.01.1481

Witt, M., Scheiner, C. W., & Robra-Bissantz, S. (2011). Gamification of online idea

competitions: insights from an explorative case.

Yanev, V. Video Game Demographics - Who Plays Games in 2022? In: Techjury.

Yee, N. (2007). Motivations for Play in Online Games. *Cyberpsychology & behavior : the impact of the Internet, multimedia and virtual reality on behavior and society*, *9*, 772-775. https://doi.org/10.1089/cpb.2006.9.772

Yee, N., Ducheneaut, N., & Nelson, L. (2012). Online gaming motivations scale: Development and validation. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/2207676.2208681

Yee, N., Ducheneaut, N., Nelson, L., & Likarish, P. (2011). *Introverted elves & conscientious gnomes: The expression of personality in World of Warcraft* (Vol. 2011). https://doi.org/10.1145/1978942.1979052

Zammitto, V. L. (2010). Gamers' personality and their gaming preferences.

Zeigler-Hill, V., & Monica, S. (2015). The HEXACO model of personality and video game preferences. *Entertainment Computing*, *11*, 21-26. https://doi.org/https://doi.org/10.1016/j.entcom.2015.08.001

Zohaib, M. (2018). Dynamic Difficulty Adjustment (DDA) in Computer Games: A Review. *Advances in Human-Computer Interaction*, *2018*, 1-12. https://doi.org/10.1155/2018/5681652

11. Appendices

Appendix A – Vignette variants for all design features

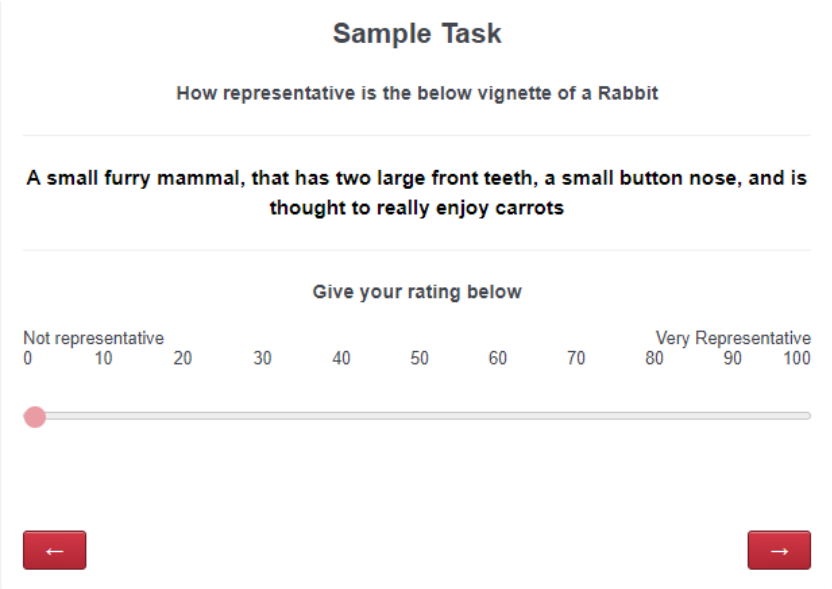
Design feature	Function	Vignette set number	Variant 1 (context excluded)	Variant 2 (context included)	Variant 3 (context included)
PVP	The element of playing against other real-life players	1	The option to compete against other players.	Being able to compete against other players	Knowing you can compete against other players
Leaderboards	Situating player progress amid the progress of other players	2	The availability to see where one's progress is situated in comparison to the progress of others	Being able to view your progress and scores in comparison to the progress of others	Being able to see how your score and rank compare with other players
Complementarity	Necessary presence of other characters when completing task (for example to complete the objective the abilities of two players must be combined)	3	The condition that requires another players' ability to help complete an objective	Being able to complete an objective only with the help and abilities of another player	Being required to complete an objective or task only with the help and abilities of another player
Shared Goal	Nonexclusive goals that can be accomplished more efficiently with other players	4	The scenario where the completion of an objective is more likely if working together with another player	Being able to complete an objective more effectively if you work with another player	Having the option to work with another player to help complete an objective or task more effectively
Trade	Transactions with other players in which advantages can be gained/shared (for example trading a rare item for a large sum of money)	5	The option to trade inventory items with players in exchange for their items, or currency	Being able to trade inventory items/currency with other players in exchange for items/currency	Being able to trade and exchange your items with other players
Text Chat	Communicating through a text channel	6	The option to communicate with other players via text chat	Being able to communicate with other players via text chat	Being able to write messages to other players
Voice Chat	Communicating through a voice channel	7	The option to communicate with other players via voice chat	Being able to communicate with other players via voice chat	Being able to talk to other players with a headset or microphone
Emotes	Communicating through avatar behaviour with the use of emotes (for example the command " /jumpingforjoy " would	8	The option to communicate with other players via avatar behaviour (such as jumping, going prone etc)	Being able to communicate with other players via avatar behaviour commands (such as jumping, going prone etc)	Being able to express your emotion and feelings through your avatar behaviour (such as jumping or dancing)

	result in 9the avatar performing a jump)				
Trophy	Evidence of merit/achievement and indication of competency	9	The scenario in which a trophy is received after completing an achievement or milestone	Being able to receive a trophy that represents your merit and/or an achievement	Being able to collect trophies after completing challenges or tasks
Badges	Evidence of merit/achievement and indication of competency	10	The scenario in which a badge is received after completing an achievement or milestone	Being able to receive a badge that represents your merit and/or an achievement	Being able to collect badges after completing challenges or tasks
Medals	Evidence of merit/achievement and indication of competency	11	The scenario in which a medal is received after completing an achievement or milestone	Being able to receive a medal that represents your merit and/or an achievement	Being able to collect medals after completing challenges or tasks
Tokens	Evidence of merit/achievement and indication of competency	12	The scenario in which redeemable tokens are received after completing an achievement or milestone	Being able to receive redeemable tokens that represent your merit and/or an achievement	Being able to earn tokens after completing challenges, that can be used to buy game content
Items	Functional objects that can be utilised to enhance skill level or abilities	13	The scenario in which redeemable items are received after completing an achievement or milestone	Being able to receive redeemable items that represents your merit and/or an achievement	Being able to receive items after completing challenges or tasks
Depletion	Sanction received following failure of some sort that reduces current inventory	14	The scenario in which a sanction is received following a failure - the sanction reduces current inventory	Having to receive a sanction following a failure - the sanction reduces current inventory (e.g., points or items)	Having to lose points, items, or currency after failing in some way
Restriction	Sanction received following failure of some sort that restricts access to game features/game area	15	The scenario in which a sanction is received following a failure - the sanction restricts access to features	Having to receive a sanction following a failure - the sanction restricts access to specific game features/areas	Having to lose access to some aspects of the game after failing in some way
Demotion	Sanction received following failure of some sort that demotes status	16	The scenario in which a sanction is received following a failure - the sanction demotes status or rank	Having to receive a sanction following a failure - the sanction demotes status or rank	Being demoted and having your rank reduced after failing in some way
Points	Numerical based indication of what the player has accrued	17	The option to view ones' progress in a numerical format	Being able to view your progress in a numerical format	Being able to see your progression in a number format
Progress Bar	Visual indication of what work is remaining before a milestone is reached/task is completed	18	The option to view a visual indication of ones' progress	Being able to view a visual indication of your progress	Being able to see how close you are to reaching a milestone in a bar format
Leaderboard	Situating player progress amid the progress of other players	19	The option to view ones' progress in comparison to the progress of others	Being able to view your progress in comparison to the progress of others	Being able to see how your score and rank compare with other players
Walkthrough	Step by step guide on any matter that will help progress through the game (for example how to perform tasks, or the rules of the game)	20	The availability of using a step-by-step guide to help learn how to play the game/perform game tasks	Being able to use a step-by-step guide that will help you learn how to play and perform game tasks	Having the option to receive a step-by-step guide on how to complete tasks or play the game

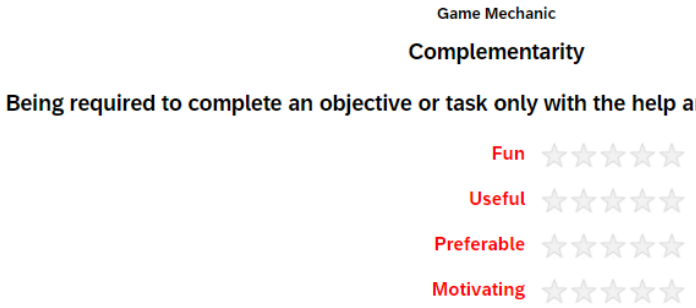
Tips / Hints	Less instructional than a walkthrough, but still provides advantageous small pieces of information that can assist with the task at hand	21	The availability to receive hints/tips when playing that will assist performing tasks and fulfilling objectives	Being able to receive hints/tips during play that will assist in performing tasks and fulfilling objectives	Being able to receive tips and hints when playing
Notification / Prompts	Reminder or notification of changes	22	The availability of receiving key notifications and updates when playing	Being able to receive key notifications and updates when playing	Being able to receive notifications while playing
Cut Scenes	Video sequences that convey story progression	23	The scenario in which video sequences can be viewed to show how the game story is progressing	Being able to view video sequences central to the plot of the games' story	Being able to see how the story progresses through cutscenes
Storyline	Context within which the game or characters are situated in	24	The story context within which every other aspect of the game is situated	Having a central theme and story that the game revolves around	Knowing the game world relates to an underlying story and lore
Currency	Accumulated spendable income	25	The feature of being able to spend earned currency within the game	Being able to spend money you have earned in the game	Being able to spend your in-game money/currency on game content
Item Degradation	The planned expiration of items possessed or purchased	26	The feature of planned expiration - where items possessed will expire at some point in time	Having the pressure of planned expiration - with items possessed having an expiry date	Knowing that you must use some items/game content sooner rather than later due to time expiration
Dashboard	Platform where game history can be accessed (e.g., resources, points, achievements etc)	27	The platform on which game history, resources, points, achievements, profile information friends list etc can be accessed	Being able to view your game history, resources, achievements, profile information, friends list etc	Being able to access game information, such as your game history, resources, profile, friends list, achievements etc
Behavioural Momentum	The game gradually increasing in difficulty	28	The gradual increase in difficulty as the game goes on	Having to play with an increasingly difficult game	Knowing that the game will become more difficult as you progress through it
Levels	Sections or parts of the game that is only accessible once a previous level is completed	29	Sections or parts of the game that are only accessible once a previous/existing level is completed	Having to continuously complete tasks and levels in order to access the next level of the game	Knowing that there are more levels to complete once the current level is completed
Barriers	Exclusion from accessing aspects/areas of the game	30	The exclusion from accessing specific aspects or parts of the game	Having to be excluded from accessing specific aspects or parts of the game	Knowing that you cannot access specific aspects or parts of the game
Game Goals	The smaller tasks and achievements to be fulfilled during play that will facilitate the game objective being completed	31	The smaller and more immediate goals that once completed, will assist in fulfilling the game objective	Having to complete smaller more immediate goals that will assist in completing the overarching objective	Knowing that there are small immediate goals to complete that will assist in completing the overarching objective
Game Objectives	The end aim to fulfil when playing	32	The overarching goal when playing a game mode	Having a overarching goal to complete when playing a game mode	Knowing there is an overarching goal to complete when playing a game mode
Design / Editing / Customisation	Opportunity to design, edit or customise aspects of the game	33	The option to edit or design aspects of the game (e.g., avatar, environment, inventory)	Being able to design or edit aspects of the game (e.g., avatar, environment, inventory)	Being able to change and design aspects of the game (e.g., avatar, environment, inventory)

Decision Making	Power to make decisions that affect the course of the game/story	34	The power to make decisions that significantly alter the course of the game/story	Being able to make decisions that significantly alter the course of the game/story	Being able to change the way in which the game story progresses, by making important/key decisions
Avatar	A virtual model/sprite/signature representation of the gamer	35	The presence of a virtual model/sprite/signature that acts as a virtual representation of the gamer	Being able to represent yourself via a virtual model/sprite/signature	Knowing you can represent yourself via a virtual model/sprite/signature
Profile	Opportunity to convey aspects of oneself to other players	36	The convention with which the gamer immediately conveys several aspects of themselves to other players	Being able to immediately convey several aspects of yourself to other players	Knowing you can immediately convey several aspects of yourself to other players
Rank / Status	The assignment of a Rank/Status to convey the players level of skill and experience	37	The category each player is assigned that reflects ability, score and/or experience	Being assigned a category and rank that reflects your ability, score and/or experience	Knowing you are assigned a category and rank based on your ability, score and/or experience

Appendix B – Visual analogue scale



Appendix C – Star rating system for design feature preference



Appendix D – PAF Correlation matrix (reduced)

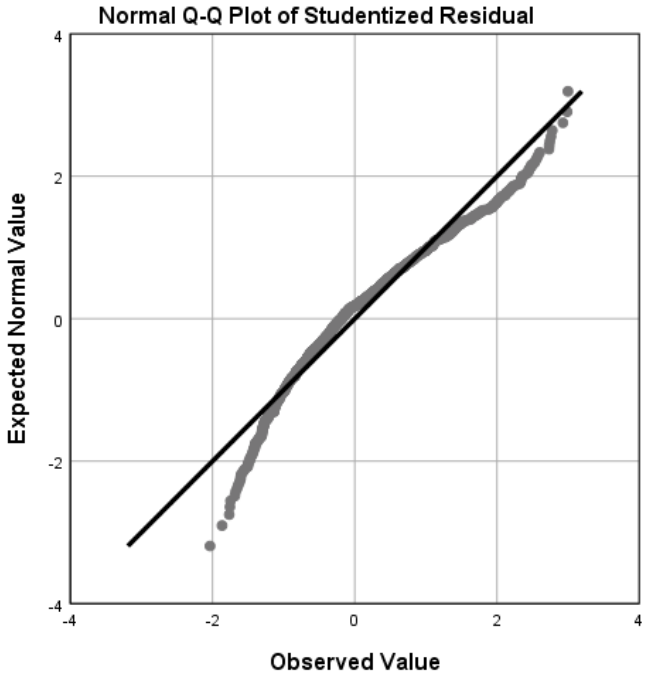
		PVP_TOTAL	LEADERBOARDCOMP_TOTAL	PUNISH_TOTAL	COMPLEMENTARITY_TOTAL	SHAREDGOAL_TOTAL	TRADE_TOTAL	FRIENDINV_TOTAL	TEXTCHAT_TOTAL
Correlation	PVP_TOTAL	1.000	.428	.310	.185	.252	.244	.363	.278
	LEADERBOARDCOMP_TOTAL	.428	1.000	.270	.187	.221	.332	.351	.290
	PUNISH_TOTAL	.310	.270	1.000	.130	.107	.169	.172	.177
	COMPLEMENTARITY_TOTAL	.185	.187	.130	1.000	.628	.132	.247	.192
	SHAREDGOAL_TOTAL	.252	.221	.107	.628	1.000	.267	.411	.252
	TRADE_TOTAL	.244	.332	.169	.132	.267	1.000	.469	.412
	FRIENDINV_TOTAL	.363	.351	.172	.247	.411	.469	1.000	.456
	TEXTCHAT_TOTAL	.278	.290	.177	.192	.252	.412	.456	1.000
	VOICECHAT_TOTAL	.342	.335	.211	.284	.374	.288	.503	.497
	EMOTES_TOTAL	.087	.119	.172	.164	.198	.294	.289	.347
	COMMUNALDISC_TOTAL	.214	.264	.108	.338	.486	.337	.391	.363
	TROPHY_TOTAL	.067	.316	.159	.085	.172	.284	.255	.181
	BADGES_TOTAL	.092	.332	.170	.127	.182	.290	.265	.216
	MEDAL_TOTAL	.099	.336	.171	.117	.144	.265	.246	.194
	POINTSTOKEN_TOTAL	.170	.254	.110	.115	.202	.353	.295	.228
	ITEMPOWERUP_TOTAL	.129	.253	.150	.144	.217	.356	.254	.156
	BONUS_TOTAL	.110	.131	.128	.126	.190	.294	.274	.219
	LOTTERY_TOTAL	.081	.099	.169	.154	.083	.149	.061	.106
	DEPLETION_TOTAL	.266	.266	.231	.195	.145	.131	.163	.136
	RESTRICTION_TOTAL	.150	.153	.226	.221	.129	.085	.065	.081
	DEMOTION_TOTAL	.337	.363	.206	.194	.175	.113	.162	.172
	POINTS_TOTAL	.203	.413	.119	.112	.212	.356	.331	.316
	PROGRESSBAR_TOTAL	.205	.374	.102	.102	.190	.352	.318	.277
	LEADERBOARDFEED_TOTAL	.396	.766	.253	.192	.259	.322	.395	.374

Appendix E – Individual KMO measures of all variables

PVP_TOTAL	.915[a]
LEADERBOARDCOMP_TOTAL	.893[a]
PUNISH_TOTAL	.904[a]
COMPLEMENTARITY_TOTAL	.806[a]
SHAREDGOAL_TOTAL	.850[a]
TRADE_TOTAL	.931[a]
FRIENDINV_TOTAL	.938[a]
TEXTCHAT_TOTAL	.938[a]
VOICECHAT_TOTAL	.937[a]
EMOTES_TOTAL	.927[a]
COMMUNALDISC_TOTAL	.947[a]
TROPHY_TOTAL	.940[a]
BADGES_TOTAL	.836[a]
MEDAL_TOTAL	.855[a]
POINTSTOKEN_TOTAL	.944[a]
ITEMPOWERUP_TOTAL	.925[a]
BONUS_TOTAL	.941[a]
LOTTERY_TOTAL	.880[a]
DEPLETION_TOTAL	.885[a]
RESTRICTION_TOTAL	.828[a]
DEMOTION_TOTAL	.911[a]
POINTS_TOTAL	.944[a]
PROGRESSBAR_TOTAL	.947[a]
LEADERBOARDFEED_TOTAL	.916[a]
SCARLETT_TOTAL	.962[a]
PERFORMANCEGRAPH_TOTAL	.966[a]
WALKTHROUGH_TOTAL	.871[a]
TIPSHINTS_TOTAL	.865[a]
NOTIFICATIONS_TOTAL	.942[a]
CUTSCENES_TOTAL	.904[a]
STORYLINE_TOTAL	.869[a]
CURRENCY_TOTAL	.933[a]
ITEMDEG_TOTAL	.921[a]
DASHBOARD_TOTAL	.974[a]
BEHAVIOURALMOM_TOTAL	.949[a]
LEVELS_TOTAL	.944[a]
BARRIERS_TOTAL	.936[a]
GAMEGOAL_TOTAL	.949[a]
GAMEOBJECTIVE_TOTAL	.937[a]
BOSSBATTLE_TOTAL	.942[a]
BEGGINNERS_TOTAL	.918[a]
DESIGN_TOTAL	.896[a]
DECISIONMAKE_TOTAL	.912[a]
VOTE_TOTAL	.960[a]
AVATAR_TOTAL	.886[a]
PROFILE_TOTAL	.930[a]
RANK_TOTAL	.954[a]

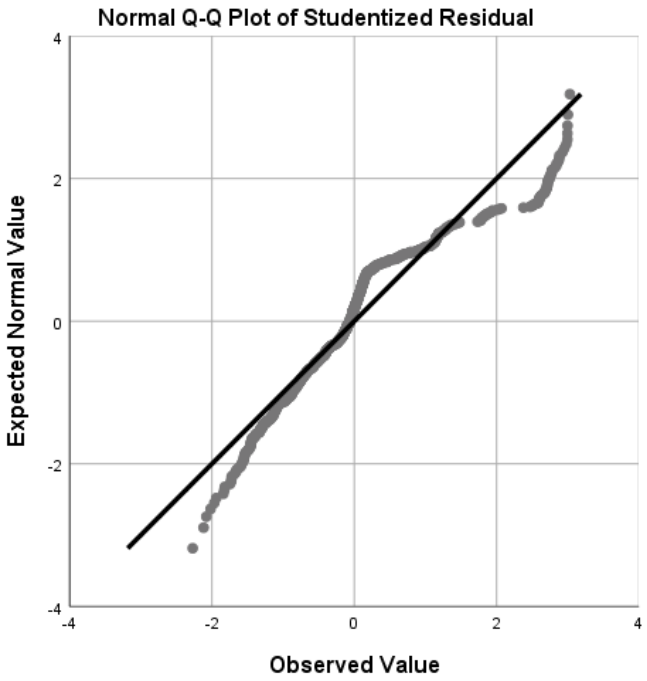
Appendix F

Q-Q plot (normal distribution) for play duration and design feature preference



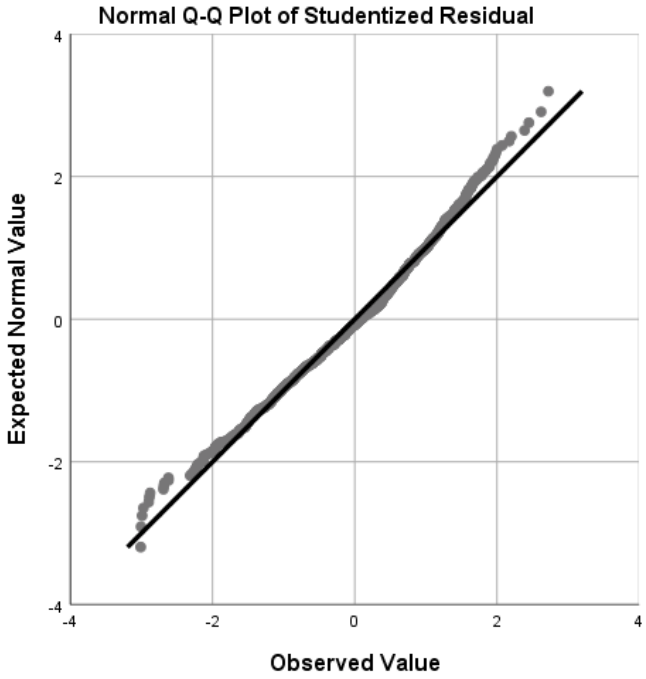
Appendix G

Q-Q plot (normal distribution) for play frequency and design feature preference



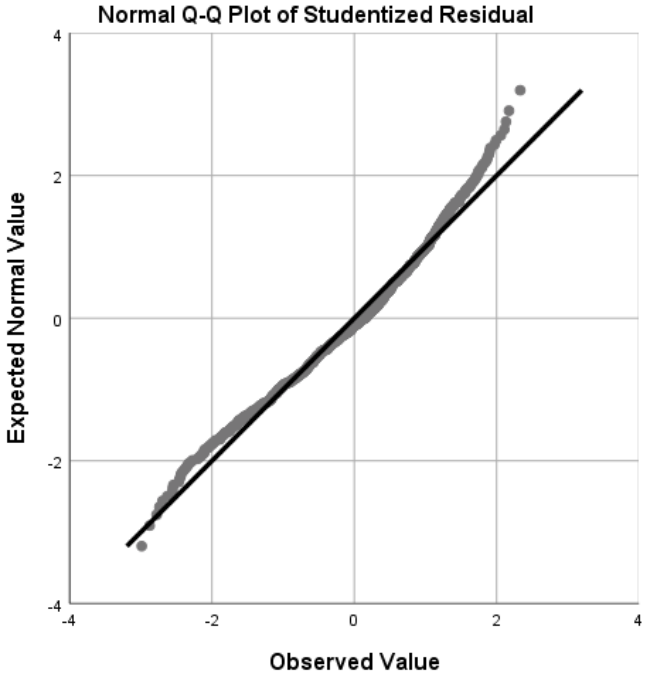
Appendix H

Q-Q plot (normal distribution) for Trojan motivations and preference for Expression design features



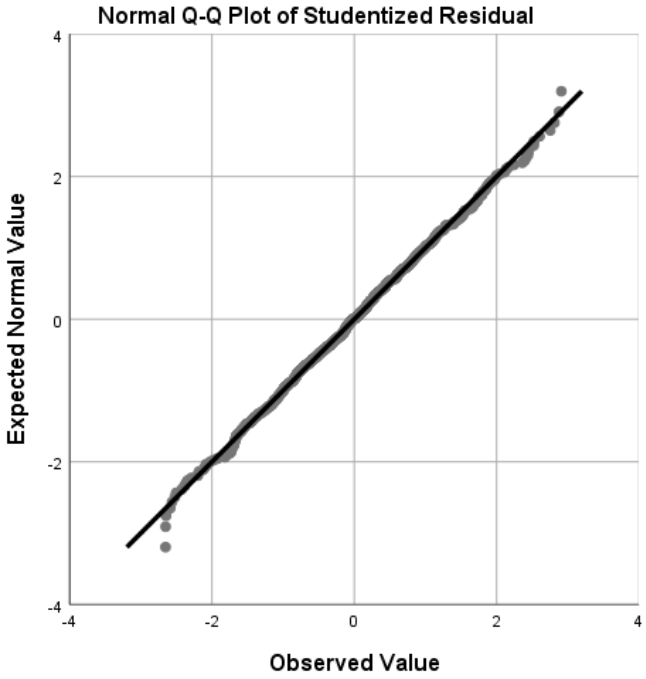
Appendix I

Q-Q plot (normal distribution) for Trojan motivations and preference for Improvement design features



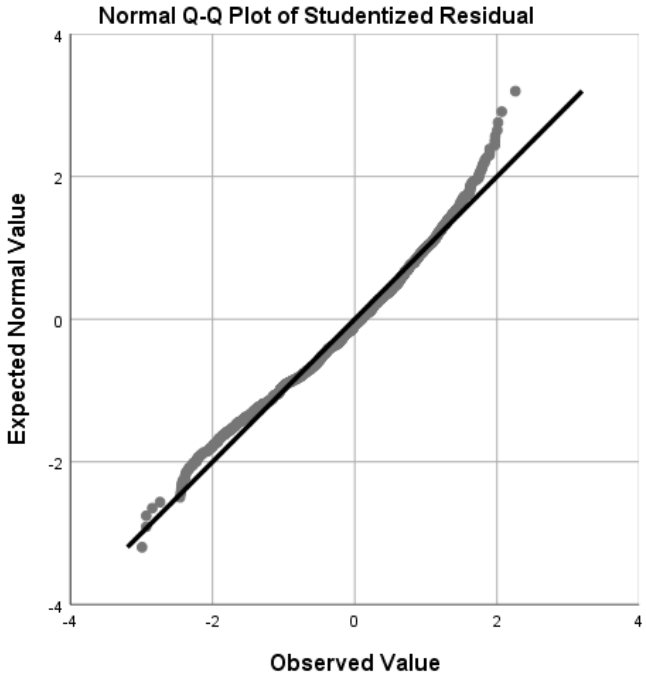
Appendix J

Q-Q plot (normal distribution) for Trojan motivations and preference for Accessibility design features



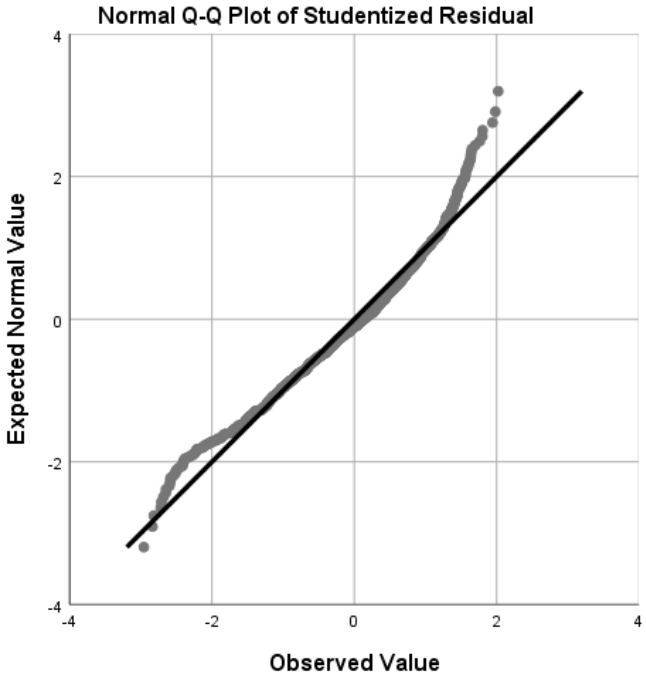
Appendix K

Q-Q plot (normal distribution) for HEXACO personality and preference for Expression design features



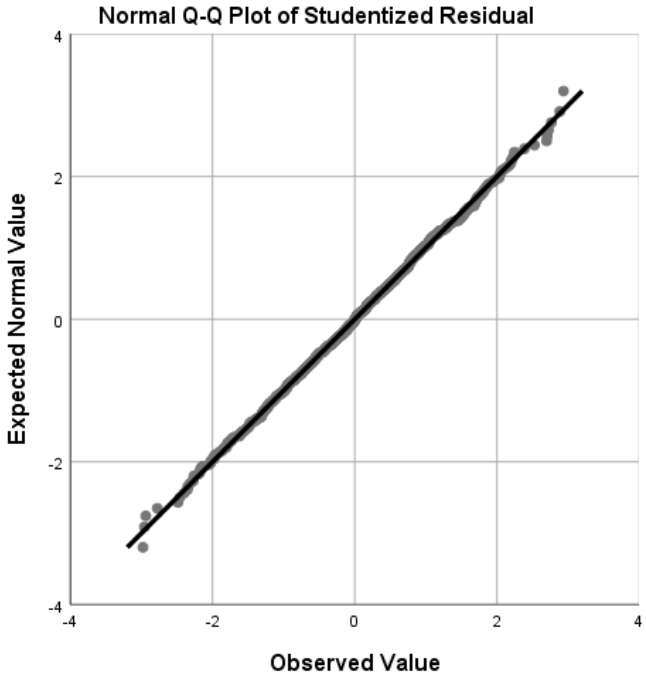
Appendix L

Q-Q plot (normal distribution) for HEXACO personality and preference for Improvement design features



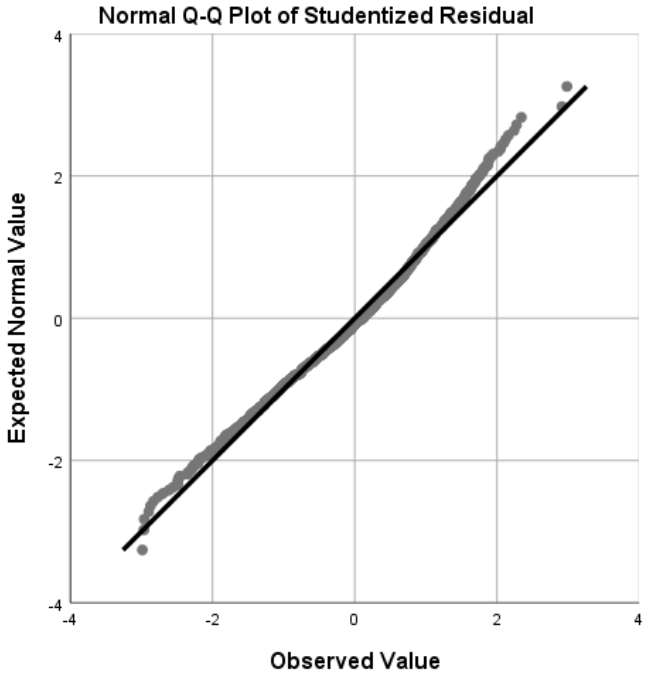
Appendix M

Q-Q plot (normal distribution) for HEXACO personality and preference for Accessibility design features



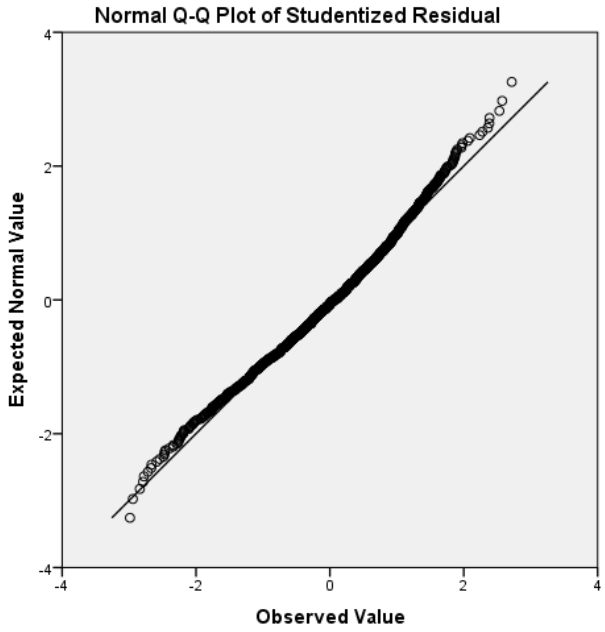
Appendix N

Q-Q plot (normal distribution) for Gamefulness and Expression design feature preference



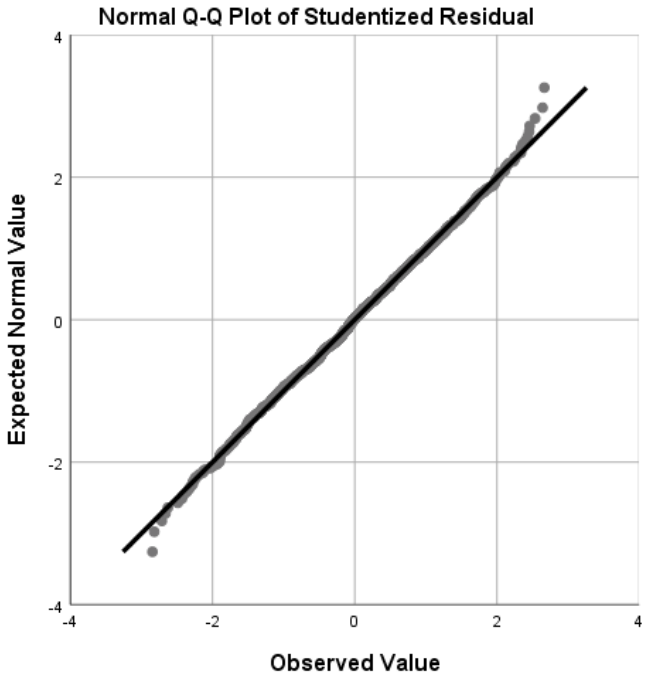
Appendix O

Q-Q plot (normal distribution) for Gamefulness and Improvement (NEW CFA) design feature preference



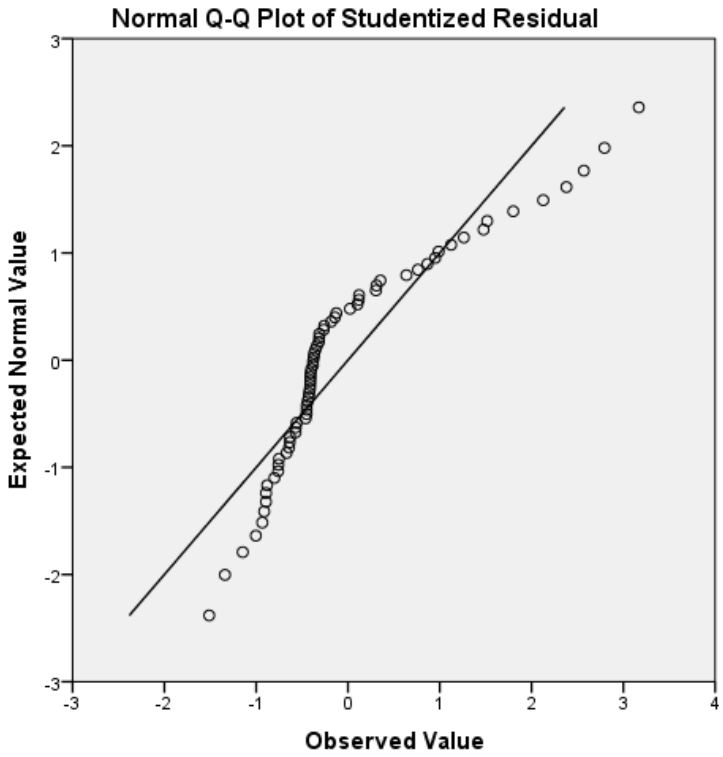
Appendix P

Q-Q plot (normal distribution) for Gamefulness and Accessibility design feature preference



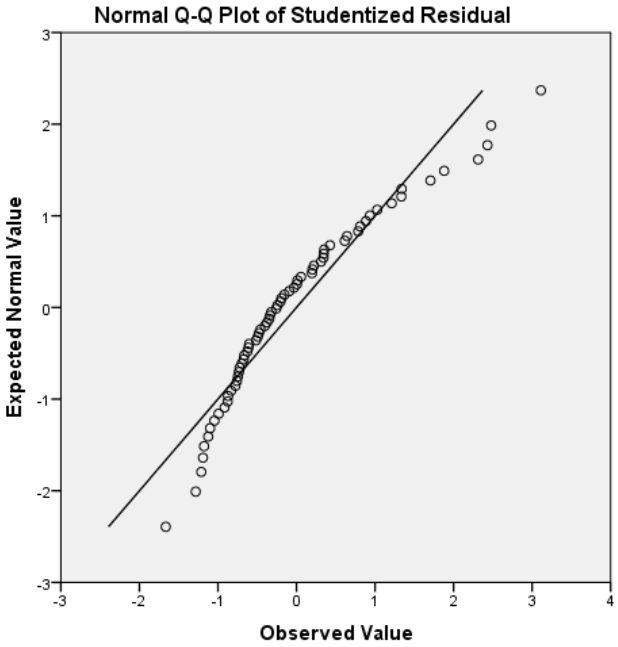
Appendix Q

Q-Q plot (normal distribution) for user design preference and engagement with an Expression game



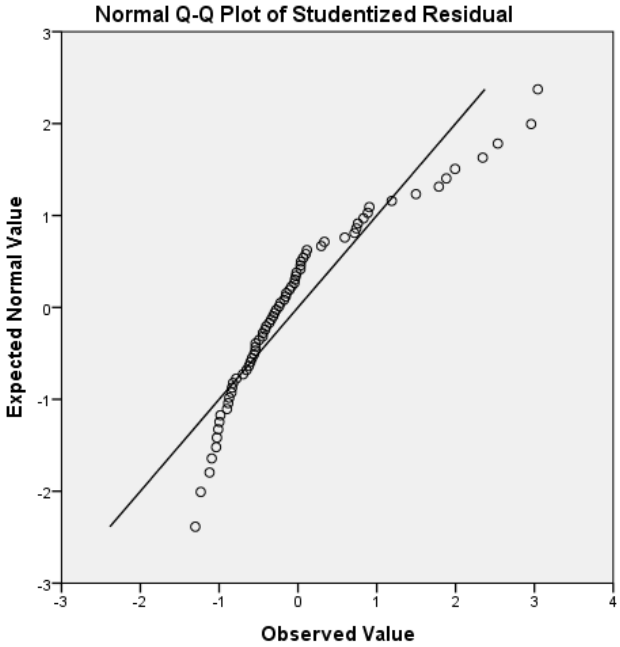
Appendix R

Q-Q plot (normal distribution) for user design preference and engagement with an Improvement game



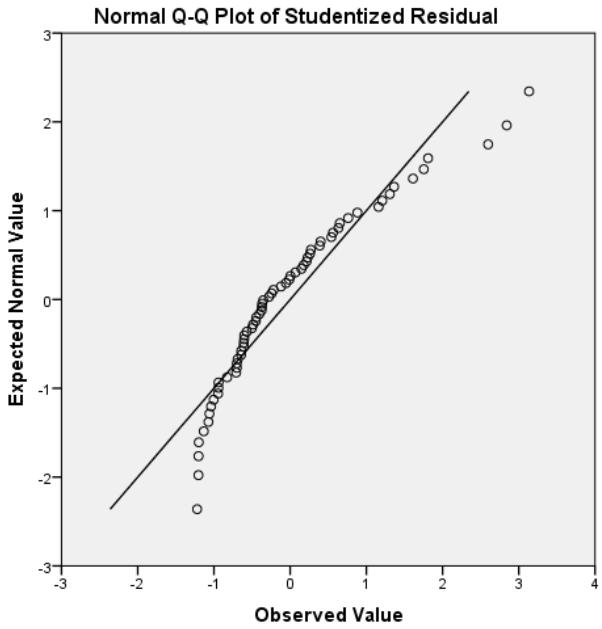
Appendix S

Q-Q plot (normal distribution) for user design preference and engagement with a Competition game



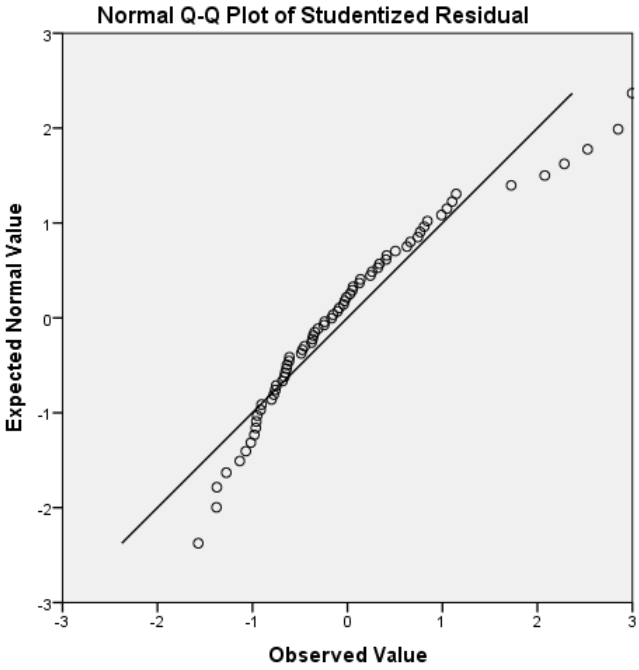
Appendix T

Q-Q plot (normal distribution) for user design preference and engagement with a Difficulty game



Appendix U

Q-Q plot (normal distribution) for user design preference and engagement with a Loss game



Appendix V

Full list of unique game titles reported as participant favourites

Assassin's Creed Revelations
Binding of Isaac
Bioshock Infinite
Chromehounds
CSGO
Death Stranding
Destiny 2
Doom 2016
Maplestory
Eliza
Fallout New Vegas
Far Cry 4
Football Manager
Grand Theft Auto V
Half-Life
Hollow Knight
Kingdom Come: Deliverance
League of legends
Life is Strange
Mafia City of Lost Heaven (or the remake)
Mass Effect
MAX PAYNE
Metal Gear Solid 3
Minecraft
Morrowind
Need For Speed Most Wanted 2005
Nier Automata
Outer Wilds
Overwatch
Path of Exile
Portal 2
Read dead redemption 2
Rome Total War
The Elder Scrolls V: Skyrim
Stardew Valley
Stellaris
Team Fortress 2
The Walking Dead
Terraria
TES III Morrowind
Tetris
The Legend of Zelda; Breath of the Wild
The Stanley Parable
The Witcher 3: The Wild Hunt
Titanfall 2
UnderTale
Warhammer
World of Warcraft

Appendix W - 47 design features

Number	Design feature	Vignette
1	PVP	Being able to compete against other players
2	Leaderboard (Competition)	Being able to see how your score and rank compare with other players
3	Punish	Being able to punish an opponent once you have defeated them
4	Complementarity	Being required to complete an objective or task only with the help and abilities of another player
5	Shared Goal	Being able to complete an objective more effectively if you work with another player
6	Trade	Being able to trade inventory items/currency with other players in exchange for items/currency
7	Friend Invite	Having the option to invite friends to play with you
8	Text Chat	Being able to communicate with other players via text chat
9	Voice Chat	Being able to communicate with other players via voice chat
10	Emotes	Being able to express your emotion and feelings through your avatar behaviour (such as jumping or dancing)
11	Communal discovery	Being able to complete tasks and learn new things as a community
12	Trophy	The scenario in which a trophy is received after completing an achievement or milestone
13	Badges	The scenario in which a badge is received after completing an achievement or milestone
14	Medal	The scenario in which a medal is received after completing an achievement or milestone
15	Tokens	Being able to earn tokens after completing challenges, that can be used to buy game content
16	Items / Power Up	Being able to receive beneficial items / power ups after completing challenges or tasks
17	Bonus	The scenario in which an unexpected or additional reward is received
18	Lottery	The scenario in which you receive a reward by luck
19	Depletion	Having to lose points, items or currency after failing in some way
20	Restriction	Having to lose access to some aspects of the game after failing in some way
21	Demotion	Being demoted and having your rank reduced after failing in some way
22	Points	Being able to see your progression in a number format
23	Progress Bar	Being able to see how close you are to reaching a milestone in a bar format
24	Leaderboard (Feedback)	Being able to see how your score and rank compare with other players
25	Scarlett letter	The element where other players are made aware of when you are stuck/failing
26	Performance graphs	Being able to view your performance level in a graph
27	Walkthrough	Having the option to receive a step-by-step guide on how to complete tasks or play the game
28	Tips / Hints	Being able to receive tips and hints when playing
29	Notification / Prompts	Being able to receive key notifications and updates when playing
30	Cut Scenes	Being able to see how the story progresses through cutscenes
31	Storyline	Having a central theme and story that the game revolves around
32	Currency	Being able to spend your in-game money/currency on game content
33	Item Degradation	Knowing that you must use some items/game content sooner rather than later due to time expiration

34	Dashboard	Being able to access game information, such as your game history, resources, profile, friends list, achievements etc
35	Behavioural Momentum	The gradual increase in difficulty as the game goes on
36	Levels	Sections or parts of the game that are only accessible once a previous/existing level is completed
37	Barriers / Access	The exclusion from accessing specific aspects or parts of the game
38	Game Goal	The smaller and more immediate goals that once completed, will assist in fulfilling the game objective
39	Game Objective	The overarching goal when playing a game mode
40	Boss Battles	Overcoming an enemy/task at the end of a level that has increased significantly in difficulty
41	Beginners luck	Help in achieving a high rate of success when completing the first few tasks
42	Design / Editing / Customisation	The option to edit or design aspects of the game (e.g., avatar, environment, inventory)
43	Decision Making	The power to make decisions that significantly alter the course of the game/story
44	Vote	Having the opportunity to vote on something (e.g., map, weapon, rules)
45	Avatar	Being able to represent yourself via a virtual model/sprite/signature
46	Profile	Being able to immediately convey several aspects of yourself to other players
47	Rank / Status	Being assigned a category and rank that reflects your ability, score and/or experience

Appendix X - Coefficients tables detailing significant and nonsignificant associations between design feature preference and engagement as measured in play duration and play frequency

Play duration and design feature dimensions

Design feature dimension	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
	B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
Difficulty	-.066	.039	-.076	-1.708	.088	.008	-.058	-.056	.551	1.814
Competition	.050	.027	.090	1.856	.064	.122	.063	.061	.464	2.153
Reward	.029	.036	.032	.815	.415	.075	.028	.027	.709	1.410
Accessibility	-.170	.037	-.178	-4.571	**.000**	-.080	-.154	-.151	.718	1.392
Loss	.015	.022	.026	0.660	.510	.054	.022	.022	.719	1.391
Expression	.094	.032	.129	2.905	**.004**	.099	.098	.096	.555	1.802
Cooperation	.009	.029	.014	.321	.748	.073	.011	.011	.591	1.691
Improvement	.099	.036	.123	2.702	**.007**	.125	.092	.089	.525	1.905
Narrative	-.084	.044	-.084	-1.917	.056	-.011	-.065	-.063	.574	1.742

Notes: bolded figures represent insignificant p values

Play frequency and design feature dimensions

Design feature dimension	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
	B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
Difficulty	-.007	.008	-.037	-.824	.410	.028	-.028	-.028	.563	1.776
Competition	.008	.006	.073	1.497	.135	.084	.051	.051	.482	2.073
Reward	-.001	.007	-.008	-.201	.841	.046	-.007	-.007	.706	1.416
Accessibility	.001	.008	.004	.098	.922	.042	.003	.003	.725	1.378
Loss	-.008	.005	-.066	-1.656	.098	-.008	-.057	-.056	.730	1.371
Expression	.015	.007	.099	2.165	**.031**	.102	.074	.074	.557	1.796
Cooperation	-.001	.006	-.009	-.203	.839	.051	-.007	-.007	.598	1.671
Improvement	.010	.008	.060	1.285	.199	.092	.044	.044	.537	1.862
Narrative	-.006	.009	-.029	-.651	.515	.035	-.022	-.022	.580	1.725

Notes: bolded figures represent significant p values

Appendix Y - Coefficients tables detailing significant and nonsignificant associations between design feature preference and user motivations

Trojan motivations (Kahn et al., 2015) and design feature dimensions

Design feature dimension	Motivation	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
Expression	Socialisers	1.135	.189	.192	5.995	**.000**	.194	.198	.184	.919	1.088
	Completionists	-.186	.251	-.025	-0.743	.458	.085	-.025	-.023	.818	1.223
	Competitors	.026	.197	.004	.132	.895	.045	.004	.004	.814	1.229
	Story-driven	1.655	.349	.181	4.748	**.000**	.309	.158	.146	.650	1.539
	Smarty-pants	.189	.331	.020	0.571	.568	.157	.019	.018	.771	1.297
	Escapists	2.312	.324	.248	7.144	**.000**	.317	.234	.219	.779	1.284
Improvement	Socialisers	1.238	.179	.230	6.921	**.000**	.254	.228	.220	.918	1.089
	Completionists	.004	.237	.001	0.017	.986	.111	.001	.001	.817	1.223
	Competitors	.292	.186	.055	1.569	.117	.124	.053	.050	.814	1.229
	Story-driven	1.255	.329	.151	3.817	**.000**	.222	.128	.121	.649	1.540
	Smarty-pants	-.010	.312	-.001	-0.032	.974	.111	-.001	-.001	.770	1.299
	Escapists	.795	.306	.094	2.598	**.010**	.140	.087	.083	.779	1.284
Accessibility	Socialisers	.305	.159	.066	1.911	.056	.065	.064	.063	.916	1.091
	Completionists	-.360	.211	-.062	-1.709	.088	.005	-.057	-.056	.818	1.223
	Competitors	.061	.166	.013	.369	.712	.022	.012	.012	.809	1.236
	Story-driven	.773	.293	.108	2.638	**.008**	.172	.089	.087	.649	1.542
	Smarty-pants	.382	.278	.051	1.374	.170	.113	.046	.045	.771	1.296
	Escapists	.832	.273	.114	3.051	**.002**	.160	.102	.100	.779	1.284

Notes: bolded figures represent significant p values

Appendix Z - Coefficients tables detailing significant and nonsignificant associations between design feature preference and user personality

HEXACO personality traits (Ashton & Lee, 2004) and design feature dimensions

Design feature dimension	Personality trait	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
Expression	Honesty-Humility	-.164	.097	-.058	-1.694	.091	-.035	-.057	-.055	.909	1.100
	Emotionality	.503	.092	.188	5.474	**.000**	.168	.181	.179	.904	1.106
	Extraversion	.155	.085	.067	1.810	.071	.054	.061	.059	.784	1.275
	Agreeableness	.257	.107	.084	2.399	**.017**	.075	.080	.078	.869	1.150
	Conscientiousness	-.045	.099	-.016	-0.452	.652	-.003	-.015	-.015	.901	1.109
	Openness	.250	.092	.092	2.705	**.007**	.117	.090	.088	.920	1.087
Improvement	Honesty-Humility	-.134	.087	-.053	-1.533	.126	-.036	-.052	-.051	.910	1.099
	Emotionality	.231	.083	.097	2.772	**.006**	.051	.093	.092	.904	1.107
	Extraversion	.244	.077	.118	3.155	**.002**	.115	.106	.105	.783	1.276
	Agreeableness	.164	.096	.061	1.702	.089	.075	.057	.057	.868	1.152
	Conscientiousness	.122	.089	.048	1.378	.169	.060	.046	.046	.901	1.110
	Openness	-.064	.083	-.026	-0.763	.446	.013	-.026	-.025	.918	1.090
Accessibility	Honesty-Humility	-.208	.077	-.093	-2.693	**.007**	-.077	-.090	-.088	.909	1.101
	Emotionality	.338	.073	.160	4.617	**.000**	.124	.153	.152	.904	1.107
	Extraversion	.160	.068	.087	2.343	**.019**	.064	.078	.077	.782	1.280
	Agreeableness	.191	.085	.079	2.243	**.025**	.065	.075	.074	.868	1.152
	Conscientiousness	.033	.079	.015	0.422	.673	.008	.014	.014	.901	1.110
	Openness	-.062	.074	-.029	-0.839	.402	.001	-.028	-.028	.917	1.090

Notes: bolded figures represent significant p values

Appendix AA - Coefficients tables detailing significant and nonsignificant associations between design feature preference and user valued Gamefulness

Gamefulness dimensions (Högberg et al., 2019) and design feature dimensions

Design feature dimension	Gamefulness dimension	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
Expression	Accomplishment	.013	.006	.099	2.046	**.041**	.220	.062	.058	.343	2.915
	Challenge	-.009	.006	-.073	-1.519	.129	.174	-.046	-.043	.343	2.912
	Competition	.010	.005	.087	2.052	**.040**	.204	.062	.058	.443	2.260
	Guided	.002	.005	.012	.324	.746	.213	.010	.009	.604	1.657
	Immersion	-.002	.004	-.016	-0.469	.639	.123	-.014	-.013	.728	1.374
	Playfulness	.018	.005	.136	3.775	**.000**	.218	.113	.107	.619	1.616
	Social- experience	.019	.003	.209	5.750	**.000**	.304	.171	.163	.609	1.643
Improvement	Accomplishment	.577	.110	.242	5.256	**.000**	.351	.157	.143	.350	2.855
	Challenge	-.266	.095	-.128	-2.792	**.005**	.263	-.084	-.076	.353	2.837
	Competition	.339	.078	.175	4.317	**.000**	.324	.129	.118	.455	2.200
	Guided	.063	.082	.027	.769	.442	.272	.023	.021	.612	1.634
	Immersion	.087	.063	.044	1.380	.168	.186	.042	.038	.740	1.352
	Playfulness	.188	.084	.077	2.247	**.025**	.211	.068	.061	.627	1.595
	Social- experience	.234	.056	.146	4.205	**.000**	.311	.126	.115	.615	1.627
Accessibility	Accomplishment	-.028	.007	-.206	-4.142	**.000**	-.206	-.124	-.120	.342	2.928
	Challenge	.018	.006	.148	2.977	**.003**	-.123	.089	.086	.341	2.931
	Competition	-.006	.005	-.055	-1.280	.201	-.156	-.039	-.037	.452	2.210
	Guided	-.003	.005	-.019	-.500	.617	-.165	-.015	-.015	.605	1.653
	Immersion	-.005	.004	-.046	-1.344	.179	-.134	-.041	-.039	.730	1.369
	Playfulness	-.007	.005	-.048	-1.309	.191	-.149	-.039	-.038	.622	1.607
	Social- experience	-.011	.003	-.121	-3.249	**.001**	-.200	-.098	-.094	.611	1.637

Notes: bolded figures represent significant p values

End