Gamification, interdependence, and the moderating effect of personality on performance

Star, Kam

Award date:
2015

Awarding institution:
Coventry University

GAMIFICATION, INTERDEPENDENCE, AND THE MODERATING EFFECT OF PERSONALITY ON PERFORMANCE

Kam Star

Serious Games Institute

Department of Computing

University of Coventry

This dissertation is submitted for the degree of Doctor of Philosophy

December 2015

To all those with a positive influence, who supported me, encouraged me, provided knowledge, insights and useful feedback, a big heartfelt thank you.

ABSTRACT

Because of their seemingly universal appeal, game elements such as points, goals and leaderboards, are increasingly being incorporated into non-entertainment situations with the aim of increasing user performance. This process is referred to as gamification. However, little empirical research exists on gamification's effectiveness in enhancing performance, particularly with respect to moderating influence of user personality traits. Social gamification that involves more than one participant incorporates social interdependence, which takes form as negative interdependence (competitive in nature) or positive interdependence (cooperative in nature). Based on the hypothesis that the interdependence type underlying a gamification system would appeal to differing personality traits, this study reports a quasi-experiment involving a platform designed to manipulate participant interdependence structure among cooperation, competition, and neutrality, with the latter acting as the control condition. These three interdependence structures functioned as the experiment's independent variable, with measures of participant performance as dependent variables, together with the participant personality traits assessed using the five factor model of personality acting as moderating variables. 294 undergraduate participants worked with the platform on a voluntary basis over an eight-week period, spending 38,180 minutes and performing 3,275 actions. At the conclusion of the experiment, the data collected were analysed using Kruskal-Wallis, ANOVAs, multilevel mixed method regression models, and a generalised estimating equation. The study's results yield significant evidence that incorporating gamification in the experimental platform increases participant performance as measured by completed actions on the platform, and that participant personality traits moderated performance depending on interdependence structure. Significant results suggest that within the gamified platform, Extraversion positively moderates performance under competition and Openness positively moderates performance under cooperation.

Keywords: gamification, interdependence, cooperation, competition, personality, performance, game elements, five factor model, multilevel mixed analysis

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my director of studies, Dr. Ian Dunwell, for his guidance, support, and patience during the whole process. Similarly I would like to thank my external supervisor, Dr. Sebastian Deterding, for his acumen and encouragement; Professor Sara de Freitas, without whom I would not have undertaken this research at the Serious Games Institute; Dr. James Brusey, who provided insight and praise; and Professor Pam Kato, who provided inspiration and support.

I would also like to offer my heartfelt thanks to Dr. Sylvester Arnab, Mr. Mike Smith, Dr. Craig Stewart, Dr. Bianca Wright, and all the students who study under their direction and who collectively were instrumental in the organisation and running of the empirical study.

I would also like to thank my family, in particular my dear wife Caroline, whose patience and support for me and our lovely children, Parisa and Roxy, continues to bring joy and fulfilment to my life. I would also like to thank my father, Dr. Kazem Memarzia, as well as my sister, Dr. Mitra Memarzia, whose achievements and thirst for knowledge are a continuous source of inspiration.

Finally I would like to thank EPSRC and PlayGen, as well as Dr. Mary Graybeal, Mr. James Allsopp, Mr. Alexander Robinson, and Mr. Gvidas Bernotas for their invaluable support.

Thank you

Kam Star

CONTENTS

LIST OF TABLES

LIST OF FIGURES

List of Appendices

1 INTRODUCTION

Driven by the desire to create competitive advantage in business, to discover new means of improving health, or to increase educational attainment, organisations and individuals from across the globe are adopting digital technologies at an ever increasing rate. In the day-to-day activities of engaging, recruiting, and training students and employees, interactive digital platforms and apps incorporating social-interaction and game elements are spearheading this adoption in efforts to increase productivity. Whether the area is education, business, health, science, culture, or policy, multiple apps or digital platforms are available, and, in fact, the number of apps being developed and consumed are growing at an exponential rate. As of July 2014, over 3 million mobile apps were downloaded over 300 billion times (STATISTICA 2014). Similarly the number of platforms offering social interaction and employing game elements and gamification has seen strong growth, with industry figures pointing to a 40% annual growth in social collaboration software (IDC 2015).

In spite of increasing digitisation, the marketing efforts of businesses developing digital learning or social collaboration platforms, and the decades of exploration by researchers in *computer-supported cooperative work* and similar fields, the major innovators of user engagement and effective performance have not been the digital learning or social collaboration platforms themselves but the e-commerce giants Amazon, eBay, and Google, which between them draw over 1 billion people a month to their platforms. Not surprisingly, the one element these organisations have in common is a serious commitment to personalisation in their user interfaces. It can be argued that humans tend to gravitate toward contexts which provide personalised solutions to their needs

and views tailored specifically to match theirs rather than to generic versions (Kramer et al. 2000). Whilst some users of these platforms are aware of the underlying recommendation and rating systems incorporated in the platform, few are aware of the deep level of personality profiling and dynamic framing performed for each individual user and of the consequent rearrangement of the platform's user interface based on the results of this profiling and framing. Amazon online retail store alone personalises over 100 million web pages daily to accommodate its visitors and drive higher sales performance (Kaptein 2015).

Adding to this digital growth is a tendency to utilise games and game elements to enhance engagement and hence productivity. Despite the hype and claims to the contrary, the majority of efforts in gamification remain generic rather than personalised (Herzig et al. 2015; Gibson and Jakl 2015), with the latter representing a potential means of delivering improved outcomes and higher performance.

This thesis explores the potential for maximising performance of users in a social collaboration platform through gamification, by examining the moderating impact of individual's personality traits on their performance under different gamification dynamics. It examines existing empirical studies in related domains to identify elements contributing to success and failure within those domains. Drawing on these findings to inform its design and implementation, the research included developing and conducting a quasi-experimental field study under controlled conditions.

1.1 The Rise of Gamification

In 2014, over a billion people played digital games on a regular basis and mostly online (Seo et al. 2015), and this rate has been increasing as more gain access to the Internet and as prices of smart device decline. Although traditionally viewed as an activity enjoyed primarily by young males, recent research suggests that roughly equal numbers of both genders enjoy digital game playing, although there are nuanced difference in the type each gender enjoys (Veltri et al. 2014, Liu et al. 2013).

The ubiquity of games and their universal appeal as a method of engagement have led to an explosion of games created for purposes other than entertainment. *Serious games* are defined as any form of interactive computer-based game software developed for purposes other than entertainment (Ritterfeld et al. 2009). Today these are used in a wide range of application domains, including learning, training, and communication, for

a range of sectors, including healthcare, culture, advertising, military, engineering, business, politics, and others, and hundreds of new serious games are being developed yearly (Laamarti 2014). *Games with a purpose* (GWAPs), related to serious games, are defined as human-based computation techniques that incorporate game elements so that the players are actually performing a computational process as they play the game (Von Ahn 2006). Such game-based techniques have been used in fields as diverse as annotating images (Seneviratne and Izquierdo 2010), collecting common-sense facts (Speer et al. 2010), and predicting protein structures (Cooper, Khatib, et al. 2010). These games are often presented as playful online toys with links back to real world data, thus enabling players to carry out such useful computational tasks as pattern recognition, to which an algorithm might be less likely to yield a valid solution. Advances in development of artificial intelligence are increasingly limiting the use of GWAPs in pattern recognition however (Cambria et al. 2015).

With games and game elements being increasingly incorporated into real-world subjects for purposes other than entertainment, it's not surprising that the Oxford English Dictionary included the word 'gamification' for the first time in 2010, a significant event testifying to the increasing incorporation of game elements into everyday activities. Work is treated as play, as every-day work or study activities are augmented with engineered systems that look like games or encompass such game elements as score points and leaderboards. Although the Oxford English Dictionary's definition explicitly identifies 'online marketing' as a typical application domain of gamification, other domains continue to be developed and explored across many other sectors including health, education, and business (Groh 2012). In the academic literature, gamification tends to be defined as the use of game elements in a non-entertainment context (Deterding et al. 2010), often with the goal of increasing engagement or productivity in a non-entertainment activity (Huotari and Hamari 2012).

Proponents of gamification argue that humans require constant feedback and acknowledgement of progress in order to carry out daily activities (McGonigal 2011). Games are prototypical examples of such feedback loops, in which granular and explicit responses to player actions are relayed back to the player, thereby enabling progress within the activity. These feedback loops are highly engaging (Munz et al. 2007), and hence borrowing specific game elements and integrating them into everyday activities is claimed to make these activities as engaging as games (Zichermann and Cunningham 2011).

Within a business world driven by increasing productivity, feedback mechanisms such as the points and leaderboards found in games are frequently viewed as robust methods for developing competition and increasing engagement (Herger 2014). Enterprise gamification is advocated as a means to relieve boredom, increase performance levels and satisfaction, and encourage unpaid work from communities internal and external to the enterprise (Mollick and Werbach 2015). Boredom has been found to destroy employee productivity and lead to higher staff turnover (Grant 2008). Games, on the other hand, are 'not boring' even though some aspects of gameplay are extremely repetitive and involve 'grinding.' Thus boredom-relieving games, some assert, can constitute a way of creating happier and more engaged workers, thereby boosting job satisfaction and improving job performance (Rothbard and Wilk 2011, Sonnentag 2012). In reality, however, whilst some games appeal to many, most games do not appeal to most people, and so incorporating game elements into everyday activity does not, as some suggest, guarantee improved productivity.

1.2 Gaps in Current Knowledge

As explored in the gamification section of the literature review chapter, research focused on gamification has increased greatly since 2010. Given the diverse application domains of gamification, this research inevitably encompasses wide ranges with respect to context of use and study designs. Therefore, it is unsurprising that, despite the large number of papers published on the subject to date, a coherent understanding of the outcomes and results of gamification is still lacking. Furthermore, the effectiveness of gamification in increasing performance in a wide range of situations remains largely unknown. As Pesce (2015) succinctly remarked on gamification, "No game works for everyone," and thus expecting any gamification to work for every member of an audience would be to adopt a one-size-fits-all mind-set, in which all people are assumed to enjoy the same experiences. Indeed, as will be discussed later, any given gamified application impacts different individuals' motivations and performances differently. Whether these differences are attributable to context or individual personality differences and personal preferences, as suggested by Hamari and Joanna (2013), constitutes a central enquiry of this thesis's empirical exploration.

The majority of gamification researchers identify points, badges and leaderboards (PBL) as the core game elements most frequently used within gamification (Danelli 2015). However, the manner in which PBL and other game elements are framed for a

user, i.e., the way in which game elements are presented to game participants through the user interface, gives rise to a variety of game dynamics (Hunicke et al. 2004). Examples of game dynamics cited from the game design domain include constraints, emotions, narrative, progression, and relationships (Feil and Scattergood 2005, Brathwaite and Schreiber 2009). This thesis focuses on the interdependence of players as presented in a gamified platform, and the impact of these relationships on performance. Measuring the behaviour of players through their interactions with a gamified platform modified to provide competitive and cooperative social interdependence through use of game elements such as points and leaderboards is intended to provide insights toward a more nuanced approach in gamification. This study will view gamification as primarily situated within a social setting and so requiring the presence of other participants in order to provide context and meaning to actions and the feedback resulting from these actions. Whilst gamification can, of course, be developed solely for individuals, the purpose of this thesis is to examine the performance effects of contrasting interdependent social dynamics within a social gamification environment.

The impact of a goals-and-reward structure on performance has been a topic attracting considerable research since the 1920s (Watson 1925). As explored in the literature review chapter, many researchers in gamification have cited two major approaches in defining desire for goal completion. One, derived from Lewin's (1935) field theory wherein an individual carries out an activity to gain the inherent satisfaction and enjoyment it imparts (Ryan and Deci 2000), is based on *intrinsic motivation*. The second, derived from behaviourist learning theory (Johnson et al. 1981) wherein an individual performs an activity for such external rewards as monetary gains, is based on *extrinsic motivation*. However whether a particular participant views a goal or reward as intrinsic or extrinsic is largely subjective and is not, as some authors have suggested, an objective truth (Richter et al. 2015). Furthermore, whilst enquiries following the extrinsic versus intrinsic delineation remain popular, the inconsistencies in participant perception call the usefulness of such a delineation into question.

As explored in detail in the social interdependence section of this thesis's literature review, social interdependence theory states that the outcome of one's actions multilaterally impacts the outcomes of other individuals' actions (Deutsch 1949, Deutsch 1962, Johnson et al. 1998, Sharan 1990). In conceptualising the social dynamics driving goal accomplishment, Deutsch developed a theory of cooperation and

competition with three types of goal structures: cooperative, competitive, and individualistic. Cooperation is generally defined as groups of individuals working together to achieve a common goal whilst sharing the outcomes, whereas competition is a zero-sum situation where individuals aim to outperform others for their own gain (Kelly and Thaibaut 1954).

The debate as to whether cooperative or competitive structures maximise motivation and performance has been the subject of considerable scientific exploration (Tauer and Harackiewicz 2004, Johnson et al. 1981). However, despite the large number of studies, definitive conclusions remain contradictory. Some theorists have argued for cooperation as the greater enhancer of productivity, particularly where a high level of interdependence exists between goals and associated tasks within a group (Harvey et al. 1961, Deutsch and Krauss 1962, Aronson and Bridgemenan 1979). On the other hand, competitive goals and reward structures have been viewed as promoting negatively valued social behaviours such as aggressiveness and hostility, in contrast to cooperation-based structures which have been thought to promote positively valued social behaviours such as good will and empathy (Ryan and Deci 1985, Barnett and Bryan 1974, Ames 1981).

Meta-analyses of studies on effects of cooperative, competitive, and individualistic goals and reward structures on academic performance support the stance that structures based on cooperation increase performance to a greater extent than do those based on individualism and competition (Johnson et al. 1999, Stanne et al. 1999). However Stanne et al. (1999) documented competition as having a greater impact on performance than cooperation depending on the manner in which the competition was presented, and Tauer and Harackiewicz (2004) report no measurable difference in performance under conditions based on competition and cooperation. Thus, in order to provide a deeper understanding of the nuanced differences between different types of competition and cooperation and to situate this study within the current literature, in section 2.2.6 of the literature review chapter a unified model of cooperation and competition that combines social interdependence theory and economic game theory is presented.

The importance of exploring the differences in performance levels within interdependent structures based on either competition or cooperation is further underlined by three factors. Firstly, the findings presented in the literature review reveal 90% (n=36) of experimental gamification studies examined rely on competition as a primary interdependence dynamic. Secondly, gamification is usually intended as a

method to enhance the performance of all individuals concerned, who are often operating within a broader group, team or class whose collective success requires cooperation within the group. Thirdly, given that researchers in the *computer-supported cooperative work* (CSCW) field have called for design of systems that reduce intra-team competition (Easterbrook 2012). And finally, since gamification researchers have called on new research to explore cooperative goals and rewards structures empirically (Danelli 2015, Butler 2015, Kotini and Tzelepi 2015).

That the dynamics underlying most gamification studies appear to rely predominantly on competition reflects the fact that competition remains the prominent dynamic in games, whether they be board games, sports, card games, or digital games. In their seminal book, *Rules of Play: Game Design Fundamentals*, Salen and Zimmerman (2004) claim that all games are inherently competitive. Even when players on the same team cooperate, there is competition between teams. Similarly even in one-person games such as Patience or Solitaire, the competition is with one's self, the object being to play the game well enough to succeed. Other authors have expressed similar views, asserting that games are predominantly competitive by nature (Juul 2011). However a growing number of researchers have called for further research into the dynamics of cooperative games (Vordere et al. 2003, Seif El-Nasr et al. 2010) and, in particular, of cooperation-based gamification (Vegt et al. 2015, Ricther et al. 2015, Nicholson 2015).

1.3 Research Questions

As discussed earlier in this chapter, despite the rapid rise in activity and the general hype within the industry surrounding gamification, questions as to the consistency of academic research containing empirical evidence on gamification's impact and outcomes with respect to performance remain (Dyer 2015, Reiners and Wood 2015). Without empirical evidence to support the validity of the methods and techniques employed in gamification, fruitful study of the subject is not possible.

Given the claimed raison d'être of gamification to positively motivate individuals so as to increase performance levels, either in terms of quality or quantity (Huotari and Hamari 2012), the first set of research questions is designed to examine this claim and is framed as follows:

Research question 1 (RQ1): Does the existing empirical evidence for gamification support the stance that it improves performance?

The following question is posed to further better understand the underlying game elements that contribute to success or failure of gamification:

Research question 2 (RQ2): What game elements are often used in gamification, and how do they impact motivation and performance?

Finally, since theory generally serves to ground and explain a phenomenon and then acts as a scaffolding on which more successful applications can be built, the following question is posed:

Research question 3 (RQ3): What prominent theories underlie design or explanation of outcomes of existing gamification experiments?

The first three research questions are explored primarily within the gamification section of the literature review chapter and rely on examination of almost 300 studies carried out between 2010 and 2015 on the subject of gamification. This study then uses the findings of these previous studies to inform the design of an empirical experiment employing the platform *StarQuest*, a social collaboration application built using an existing open source blogging platform, which provides a private online environment for small groups of individuals to find and share digital content.

The literature review chapter will demonstrate that as many as 84% of gamification studies (n=55) result in either no improvement or a reduction in performance. Therefore, the next research question concerns the impact on performance of incorporating gamification into *StarQuest* through a quasi-between-subject experiment:

Research question 4 (RQ4): What is the impact of incorporating gamification on participants' performance within *StarQuest*?

Given the inconsistent findings regarding effects of competitive and cooperative interdependence dynamics on performance and the lack of studies on the impact of these dynamics incorporated within gamification, the following question concerns manipulation of the framing used in communicating such game elements as points, goals, leaderboards, and status in *StarQuest*:

Research question 5 (RQ5): Which gamification dynamic, competitive or cooperative, results in higher participant performance in *StarQuest*?

Finally, as indicated in the findings reported in the literature review, individual differences in participant personalities constitute a possible predictive factor in the impact of gamification on performance. Personality traits refer to the enduring personal characteristics that a particular pattern of behaviour exhibited in a variety of situations reveals. One of the most widely utilised models of personality is the Five Factor Model (FFM) proposed in the field of trait theory by McCrae and Costa (1987). The five traits that compose the FFM are Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism, and empirical research on these traits shows consistency in measurement across interviews, self-reports, and observations, across a wide range of individuals of differing ages and cultures (Schacter et al. 2009, Barrick et al. 1993).

Measurements of the five FFM personality traits have been shown as moderating factors with respect to performance in a wide variety of activities that include commercial work oriented activities (Barrick et al. 1993, Colbert et al. 2004); learning and academic performance (Furnham et al. 1999, Yeung et al. 2012, Noftle and Robins 2007, O'Connor and Paunonen 2007); Internet usage (Landers and Lounsbury 2006); and play (Teng 2008, Yee 2006). The link between personality traits and performance in discrete activities across learning, work, and play has been well established; however, still unknown are the moderating effects of personality on performance levels when game elements are incorporated into learning and work activities, i.e., gamification. Thus the sixth research question is as follows:

Research question 6 (RQ6): What moderating effects do personality traits have on relative performance levels under the differing gamification dynamics, competition or cooperation, in *StarQuest*?

This thesis presents empirical measurements of individuals' performance levels in the *StarQuest* online platform manipulated so as to operate under conditions that represent cooperation, competition, and neutral environments. With respect to performance, many factors beyond the control of the experimenter can impact the measured results. Among these are environmental, personal, cultural, interpersonal (e.g., how long participants have known one another, whether they like one another or not), situational, or contextual (e.g., attitude toward their organisation) ones. Other potential influences are the quality of the user interface and users' familiarity with and acceptance of an experiment's platform(s). Since the experimental study is carried out in-the-field rather than under laboratory conditions, ignoring these real-life factors could result in falling for the "the ludic fallacy," a term coined by Taleb (2007). When seeking to describe the world in terms of mathematical models and applying simplistic statistical models in complex domains where incorporating the entirety of the available knowledge in the model is impossible, small unknown variations in context can substantially impact outcomes (Judge et al. 2014, Steffensmeier 2008, Spence and Keeping 2010).

1.4 Contribution of this Thesis

This thesis makes at least four original contributions to the burgeoning areas of gamification, particularly online social gamification, and of related research and application fields.

Firstly, the analysis involved in critically examining the existing literature and empirical studies on gamification provides the evidence to debunk the myth that gamification is a universal solution to increased performance, as industry and some researchers in the field contend. Moreover, it provides a measured approach to identifying the conditions under which it succeeds or fails.

Secondly, by contrasting the views of social interdependence presented in social psychology literature and studies and in economic game theory, it provides a unified approach to examining cooperation and competition and so lays the foundation for a framework incorporating the two that extends and unites the current models used to describe interpersonal interdependencies.

Thirdly, it provides empirical evidence aimed at determining whether gamification is an effective method of increasing performance, extending this exploration by drawing on research on differences based on personality traits to provide evidence on the effect these traits have on performance level in a gamification context.

Finally, to the best of this researcher's knowledge, the study described in this thesis provides the first evidence of the impact on performance level of personalised gamification. Platforms able to adapt according to an individual's personality traits are customisable at the social dynamic level and so represent a move away from the one-size-fits-all approach. As the results of this study demonstrate, customisable gamification has the potential to maximise individual performance.

Whilst the research reported in a PhD thesis is primarily aimed at extending existing knowledge and exploring insights that have not yet been investigated, the ambition motivating this research has been to not merely discover new knowledge for its own sake but rather to develop practical applications that will fundamentally improve performance, particularly with respect to groups collaborating online to accomplish a task. The design and development of this research's experimental platform *StarQuest* has produced a practical social collaboration platform capable of delivering personalised game dynamics that is being used by hundreds of students daily and that continues to grow and develop.

1.5 Thesis Structure

The thesis is composed of five chapters: the introduction, literature review, methodology, results, and conclusions. The first section of the literature review chapter provides a detailed exploration of the current state of the art in empirical studies on gamification together with respective theoretical groundings aimed at answering the first three research questions. The literature review's remaining sections examine social interdependence and competition and cooperation theories, discuss personality and the impact of personality on performance, and synthesize interdependence, personality, and performance in order to develop the hypotheses for the empirical experiment to test. The methodology chapter justifies the research philosophy, paradigm, and strategies by describing the methods used for designing the experimental instrument and discussing the operationalisation of the variables and their validity. It also discusses the study's participants, analyses, and limitations. The results chapter provides descriptive and inferential analysis of the data generated from the empirical study in support of the hypothesis presented. Finally, the conclusion chapter sums up the findings presented throughout the thesis, discusses the study's limitations, and describes future directions for research and development.

1.6 Summary

Gamification appears to be a new set of tools to provide ever deeper levels of engagement and increased productivity. Despite its apparent promise, however, not every individual likes every game, and no definitive answer as to why gamification works or not yet exists. Whilst the majority of gamification applications have drawn on the competitive nature of games, much of learning and working is collaborative in nature. Introducing competition into environments that are inherently collaborative can produce negative effects, but is collaborative gamification the game that everyone wants to play? And given that users of any system are unlikely to have homogenous personalities, do some individuals perform better in a cooperative environment than in a competitive one and vice versa? The remainder of this thesis will explore the potential to increase productivity through better understanding of how personality may moderate different types of social gamification.

2 LITERATURE REVIEW

This chapter provides a comprehensive review of literature pertaining to the relevant topics explored in this thesis. It is divided into four sections. In the first section, gamification and the current state of the art with respect to empirical studies together with their theoretical groundings are explored in view of the first three research questions. In the second section, social interdependence and alternative perspectives on competition and cooperation theories are explored leading to presentation of a unified framework for competition and cooperation. In the third section, personality traits as relating to games and gamification studies and the impact of personality on performance are investigated. Finally, in the last section, research literature spanning interdependence, personality, and performance are synthesised and developed into key hypotheses to be tested in the empirical experiment.

2.1 Gamification

2.1.1 Situating Gamification

The goal of gamification is to increase participant engagement and influence participant behaviour by incorporating game elements in non-entertainment activities (Deterding et al. 2011). As shown by the academic literature, gamification is present in many disciplines, including computer science, social science, business and management, psychology, decision sciences, and economics. Within computer science, gamification research is commonly situated within the areas dedicated to human factors, human-computer interaction, and computer-supported cooperative work and social computing.

Incorporating gaming or gaming elements in non-gaming activities is not a unique definition of gamification. Different researchers employ the word 'gamification' to denote subtly different concepts, including serious games, i.e., ones created to achieve a non-entertainment purpose; augmentation or enhancement of an existing system or activity using game elements. Table 2.1 below provides a short summary of related terminologies.

Table 2.1: Related Terminologies to 'Gamification.'

Concept	Definitions	Goals
Gamification	Use of game elements in a non-entertainment context (Deterding et al. 2011). Process of enhancing a service with affordances for gameful experiences in order to support the user's overall value creation (Huotari and Hamari 2012).	Increase engagement and value creation in a non-entertainment activity
Serious Games	A mental contest, played with a computer in accordance with specific rules, that uses entertainment to further government or corporate training, education, health, public policy, or strategic communication objectives (Zyda 2005). Any form of interactive computer-based game software for one or multiple players for use on any platform developed to be more than entertainment (Ritterfeld et al. 2009).	Accomplish training, change behaviour, raise awareness across a diverse set of sectors including scientific exploration, healthcare, engineering, business, politics, culture, etc.
Games with a Purpose	A human-based computation technique in which humans perform a computational process through incorporation of game elements (Von Ahn 2003).	Carry out useful human computational tasks as a side benefit of game playing.
Edutainment	A combination of education and entertainment, refers to any entertainment content that is designed to educate as well as entertain (Egenfeldt-Nielsen 2005).	Educate whilst also entertaining.
Productivity Games	Games played to achieve some productive, useful outcome which may not necessarily be computation based, i.e., but community improvement (Smith 2011).	Create useful outcomes or increase productivity of a task through use of games.
Persuasive Technology	Interactive technology designed to change user attitudes or behaviours through persuasion and social influence	Cause user behaviour modification or attitude change.

	but not through coercion (Fogg 2002, Oinas-Kukkonen et al. 2009).	
Games	Structured play, usually undertaken for enjoyment. Games are distinct from work, which is usually carried out for remuneration, and from art, which is primarily an aesthetic or ideological expression (Wittgenstein et al. 1958). A game is a system in which players engage in an artificial conflict defined by rules and resulting in a quantifiable outcome (Salen et al. 2004).	Uses goals, rules, challenges, and interaction to provide fun by producing uncertain, non-productive, fictitious outcomes (Caillois and Barash 1961).
Choice Architecture	Design of the manner in which user decisions can be presented and framed so as to influence people's decisions (Thaler and Sunstein 2008).	Aid people in making specific decisions.

The concepts embodied in the terms *serious games*, *game-based-learning* (GBL), and *games with a purpose* (GWAPs) appear most closely related to gamification. However, while in these games, thinking is applied to real world problems, the outcomes are usually actual games, acting either as safe environments in which to practice or learn new skills or as methods to provide 'fun' whilst the players carry out useful computational tasks. In contrast, gamification, as described by Deterding et al. (2011), is the incorporation of game elements in real-world systems as a means of increasing motivational affordances of the system and enhancing engagement.

To understand the differences in these concepts, consider 'improving driving ability' as an example. A simulation of driving based on a computer or other non-automobile device and providing context and guidance would constitute a serious game or game-based learning approach to driving. A gamification approach, on the other hand, would incorporate elements designed to improve driving abilities within an actual vehicle and provide user feedback, for instance how economical a particular driving style is, as the user drives. A GWAP approach could have the driver report traffic conditions to a central system to improve route planning for others. Distinguishing between whether an experience is a game or another type of activity augmented with game elements could be subjective (Richter et al. 2015).

Recent economic models of human performance incorporating the social and behavioural dimensions of activities have viewed individuals as driven by social influences, specifically a tendency to copy the behaviour of others and a desire to fit in (Earls 2015). Dubbed behavioural economics (Jolls et al. 1998), these models are being

increasingly used in gamification studies (Butler 2015). Choice architecture, framing the message so as to influence user decision making (Tversky and Kahneman 1981, 1986), employs cognitive biases to increase the emotional appeal of one choice relative to that of another. Social games such as Farmville, and many others played on social networks, repurpose such concepts as reciprocity, i.e., the desire to give back what one has received from others, into game elements based on gift giving and designed to sustain player engagement in the platform (Deterding 2015). Also aimed at ensuring long-term engagement (Hamari 2013) are such game elements as *resource decay* and *endowment effect*, where artificial reduction in resources coupled with a sense of ownership increases participant persistence in attempting to reach a goal (Nunes and Dreze 2006). *Appointment* game elements nudge participants to return to play the game at predetermined times. For example, in a randomized controlled trial to improve clinicians' ability to measure blood pressure, gamification using appointment game elements has been found to significantly reduce the time required to take measurements (Kerfoot et al. 2014).

The boundary between persuasive technology and choice architecture within gamification is not clear cut, and gamification could be viewed as a subset of the former in so far as most examples of gamification use interactive technologies to influence users to perform a particular activity. Choice architecture, on the other hand, approaches decision points within an activity through the lens of behavioural economics, paying particular attention to the effects of framing, default selections, and perceived social norms (Butler 2015). Players' actions and the associated rewards in gamification can be viewed as directed and created to accomplish goals. The use of PBL game elements, in particular the framing of points and leaderboards, can be manipulated to create different reward structures and so affect the means and outcomes of a gamified system.

Opponents of gamification label it as 'exploitation-ware' and argue that it extracts real value from users and employees in return for mere virtual tokens (Bogost 2011). Gamification described thusly is a strategy to captivate attention and propel participants toward purchasing commodities or pursuing activities that generate revenue and, moreover, it does so by altering the way the activity or commodity is represented rather than fundamentally changing its underlying purpose. Reassigning new signifiers to the activity and redefining the activity's perceived purpose (Rey 2015) makes gamification a type of *playbour*, a term coined by Kucklich (2005) to denote activities that combine play and labour and so ensnare workers through their need to earn a living and their

commitment to self-expression. Whilst play may be voluntary, forced play in combination with labour, as in some gamification applications, is viewed as leading to exploitation, which presenting 'work' as play makes more efficient by minimising worker resistance and by increasing worker interest and motivation. In effect, it is argued that gamification exploits participants by not offering a genuine or meaningful reward but rather by offering virtual rewards that are designed to increase the profits of those who designed the system (Rey 2015).

Some challenge the validity of the term *gamification*, asserting that combining the word *game*, a complex and not easily understood medium, with the *-ification* suffix, which denotes a process of becoming, creates a word analogous to beautification (a process or mechanism that aims to renders the unsightly attractive) and falsification (a process or mechanism that aims to change something in order to make people believe something that is not true) (Bogost 2015). However it can be argued that gamification, like any other form of persuasion or cognitive manipulation, embodies nothing inherently evil or deceptive, but, much like any other medium or technology, can be used positively or exploitatively (Anderson and Rainie 2012).

Based on the gamification literature, the primary purpose underlying studies on gamification is to increase engagement and performance, although these studies fail to form a coherent discourse on the topic. Moreover, the evidence presented in support of gamification often fails to include a thorough exploration of its effects on performance. Hence a thorough survey of the empirical evidence so far provided for gamification was necessary in order to identify leading research inquiries and current gaps in knowledge. The next section offers a detailed literature review on gamification, with a particular focus on empirical evidence and theoretical foundations.

2.1.2 Empirical Studies on Gamification

Gamification as a whole is a relatively new field of study. The first studies explicitly using the terminology appeared in late 2010, and thus far very few established theoretical frameworks or a unified discourse has evolved (Landers et al. 2015). As of the beginning of 2015, only two meta-analyses of empirical studies on gamification have been published.

The first, carried out by Hamari et al. (2014), examined 24 peer-reviewed studies focused on the motivational affordance of gamified systems, i.e., the gamification

elements of a system's propensity to fulfil users' motivational needs. The review indicated that gamification seems to produce positive effects, although it does not distinguish between effects on motivation and those on performance and it cautions that outcomes are greatly dependent on context as well as user (Hamari et al. 2014).

More recently, Seaborn and Fels (2015) carried out a systematic survey on the use of gamification from a human-computer studies perspective, noting the diverse meaning and contradictory uses of the term gamification, the little empirical work performed to validate gamification concepts, and incongruities within the empirical findings.

Neither meta-analysis revealed whether gamification is effective in increasing performance. Given the current limited knowledge within the domain, critically examining the efficacy of gamification as it impacts performance through a detailed and systematic literature review following the methodologies of Cook et al. (1997) and Tranfield et al. (2003) was undertaken.

Research questions Q1 to Q3 as highlighted in the introduction served as a starting point for this initial literature review, which focused on the sum of empirical evidence of gamification's impact on user motivation and performance, the use of game elements within gamification, and the theoretical frameworks utilised within the design or explanation of the study results.

2.1.2.1 Data Collection and Analysis Method

The terms 'gamification' and the wildcard 'gamif*', which encompasses such alternative verbs to 'gamification' as 'gamifying', 'gamify', 'gamified', and 'gamifiable', in study body and title searched in *Google Scholar*, *Microsoft Academic Research*, *ScienceDirect*, *ACM Digital library*, *EBSCOHost*, *IEEE*, *Elsevier*, *Ethos*, and *Scopus* yielded a total of 2750 published papers. Filtering these studies to ones specifically containing the keyword in either title or abstract produced a list containing around 300 studies.

Those studies from among the 300 containing empirical research were identified. Peer-reviewed empirical studies containing reports of findings on gamification impacts and outcomes using only randomized control trials (RCTs) and quasi-experiments constituted a relatively small number of studies ($n < 20$). For this reason, the type of studies to be reviewed was broadened to include case studies and surveys, thus

increasing the total number of studies for detailed examination to 55. Appendix 1 contains the list of coded studies together with a summary of high level outcomes.

Due to the widely diverse reporting used in these studies, computing effect sizes for a systematic meta-analysis as described by Borenstein et al. (2009) did not provide a comprehensive outcome. Therefore a literature review based on a similar review by Connolly et al. (2012) on empirical evidence of serious games was undertaken, and effect sizes taken into consideration where available.

These studies (n=55) were coded using a data extraction pro forma (Connolly et al. 2012) and categorised as follows according to several salient dimensions:

> *Impact of gamification on participants' motivation*: Examining the observed impact of gamification on participants' motivation. Exploring theoretical frameworks used to underline or explain the effects.
>
> *Impact of gamification on participants' performance*: Examining the observed impact of gamification on participants' performance in terms of quantity or quality of output or, where available, on accuracy and speeds.
>
> *Use of game elements in gamification*: Investigating the use and effects of specific game mechanisms, also referred to as motivational affordance mechanisms, on outcomes.
>
> *Theoretical frameworks used in gamification*: Considering underlying theories, identified through citations and references made by each study's authors.

The following sections report the empirical observations made from the studies explored.

2.1.2.2 Impact on Participants' Motivation

Motivation is a theoretical construct that encompasses internal and external factors which cause an individual to behave in a specific way (Padree 1990). Its study seeks to explain causality, intensity, and persistence of behaviour. The general theme of motivational theories is to integrate and unify assumptions into a coherent whole, including the notions that motivation benefits adaptation, directs attention, influences the stream of behaviour, includes approach and avoidance tendencies, and reveals wants and desires (Reeve 2001).

Increasing user motivation is regarded as a key defining feature of gamification (Zichermann and Linder 2010).

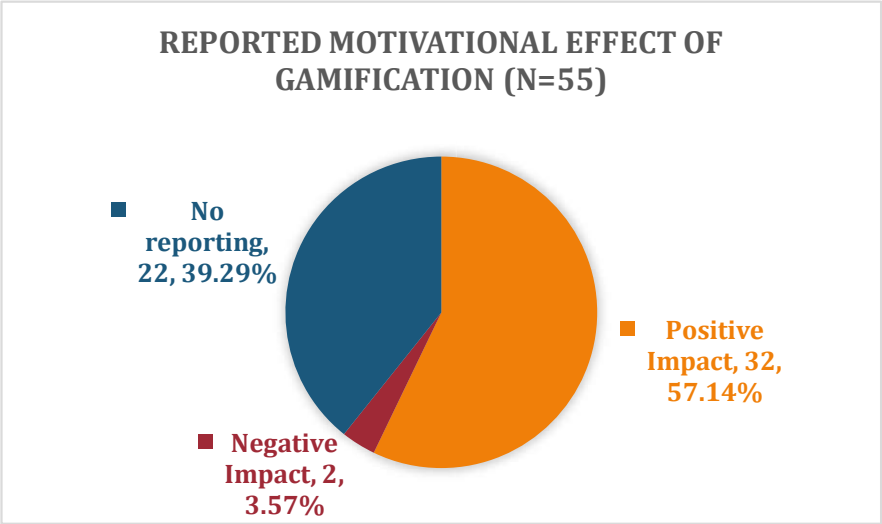


Figure 2.1: Motivational Effects of Gamification.

From the 55 gamification studies examined, 62% (n=34) reported on the motivational aspects of participants engaging through gamification. Of these studies, 94% (n=32) reported a positive motivational effect resulting from incorporation of gamification, strongly suggesting that gamification consistently increased participant motivation in completing or undertaking study tasks. This finding is consistent with the general findings of the gamification meta-analysis of Hamari et al. (2014).

Unfortunately, the methodologies used to measure motivation across studies were inconsistent. Just two studies (n=55) used a validated questionnaire; Ong (2013) assessed motivation through two sub-scales of the Dundee Stress State Questionnaire for Success Motivation and Intrinsic Motivation (Matthews et al. 1999 and 2002); and Brauner (2013) measured motivation by a scale from Schuler (2004). Only one study considered the longitudinal impact of motivation (Hamari et al., 2013), and 40% (n=22) of the studies examined did not report any motivational factors. Whilst the decision to exclude motivation may be due to an alternative focus on the part of the researchers, the possibility that the researchers were reluctant to report negative results, a form of publication bias established by Olson et al. (2002).

Critically, whilst existing research and literature on the link between motivation and performance reports an unequivocal positive correlation, when factors such as ability and opportunity (Blumberg and Pringle 1982, Siemsen et al. 2008, Wyatt 1934) are controlled for, a positive correlation between motivation and performance was not

reported. Thus, the reports on motivation, as indicated by the gamification studies explored, do not appear to provide a reliable measure of outcomes.

2.1.2.3 Impact on Participants' Performance

Gamification efforts are generally aimed at improving engagement in order to increase the performance or outcomes associated with a given situation, whether this be achieving higher levels of student learning, increasing sales, improving collaboration between colleagues, or increasing visitor effort level. In summary, the underlying purpose for incorporating gamification is to encourage participants to expend more energy and perform at higher levels (Unger et al. 2013).

Empirical studies on gamification typically report impact or changes in performance or productivity, i.e., effectiveness of effort, due to the introduction of gamification elements. Ideally, such studies incorporate a control or pre-post setup in comparison to which the effects (with respect to quality or quantity of outcomes or participant's productivity) of incorporating such elements can be measured. Over half, 54% (n=30), of the studies incorporated such a control or pre-setup measure. The failure to measure output impact in 44% (n=24) of the empirical studies, certainly a matter of concern, was partially due to the inclusion of case studies and post-hoc questionnaires or other qualitative methods which lacked an effective means of comparison with non-use of gamification.

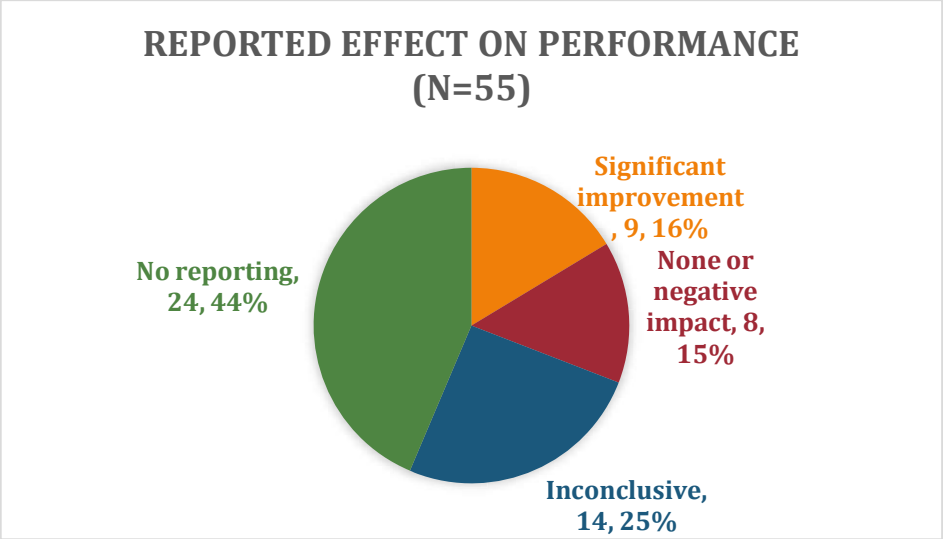


Figure 2.2: Reported Effect on Performance.

Although 25% (n=14) of studies claimed to report gamification's effect on performance, these provided no conclusive or comparative measurements. Additionally, a further

15% (n=8) of studies reported no statistically significant impact or decrease in performance due to its presence. Thus, taken together, in 84% of empirical experiments on gamification, impact on performance is either not reported, inconclusive, statistically insignificant, or negative. Put another way, in just 16% (n=8 from n=55) of all studies examined, output was reported as significantly improved, a finding that agrees with the meta-analysis by Hamari et al. (2014), wherein 10% (n=2 from n=22) of studies examined suggested a significant positive effect across measured factors.

Further validating the finding that 84% of studies fail to support gamification as a method of increasing performance is research institute Gartner's prediction in November 2012 that, by 2014, 80% of gamification attempts would fail to increase performance. In a private, unpublished interview, Brian Burke (MWC2015) explained the methodology used to derive the figure 80%. The total number of commercial initiatives which publicly announced their intention to use gamification was later compared to the proportion of those subsequently reporting positive impacts, and only around 20% of projects reported positive outcomes. The Gartner report therefore inferred that, for the remaining 80%, gamification had probably failed to live up to expectations, most likely due to inappropriate design.

In summary, gamification appears to be less effective in boosting performance than much of the marketing and hype (Zichermann and Cunningham 2010) would suggest, thus the following sections explore the elements that contribute to its success and failure.

2.1.2.4 Theoretical Frameworks in Gamification

Theory can serve to ground and explain a phenomenon, and a good theory can act as a scaffolding for building more successful applications. From the 55 empirical studies examined, 60% (n=33) cited at least one theoretical framework as a part of the experimental design or an explanation of observed outcomes. From the studies citing theoretical framework, 39% cited self-determination theory (Ryan and Deci 1985), 27% cited flow (Csikszentmihalyi 1991), 24% cited intrinsically motivating instruction theory (Malone 1981), 18% cited player type (Bartle 1996), 15% cited goal-setting theory (Locke et al. 1990), and 12% cited hierarchy of needs (Maslow 1943). Other theories mentioned by less than 10% of the sub-sample included social proof (Cialdini 1993), game flow (Sweetser and Wyeth 2005), cognitive dissonance (Festinger 1962),

expectancy x value (Vroom 1964), self-efficacy (Bandura 1997), and achievement motivation (Elliot and Church 1997).

Reeve (2001) provides a comprehensive and detailed analysis of current motivational theories, incorporating 24 theories in an overarching framework that explains motivation in terms of behaviour and physiology. Of these 24 primary motivation theories (Reeve 2001), only four theories were cited in the gamification studies examined for this study; these four—self-determination (Ryan and Deci 1985), flow (Csikszentmihalyi 1991), hierarchy of needs (Maslow 1943), and goal setting (Locke et al. 1990)—are commonly used by researchers in the gamification field. This finding indicates that gamification researchers have yet to reconcile their efforts with the established field of motivational research, but recent researches, such as that by Richter et al. (2015), have proposed unified models of motivation with respect to gamification.

The next section explores researchers' use of the most cited motivational theories, specifically, their application in defining or explaining the phenomenon of gamification and its impact on performance.

2.1.2.5 Motivation and Performance

The observed disparity between reported motivational effects and performance in the gamification literature requires further exploration. As noted by Seaborn and Fels (2015), some authors within the field appear to lack an in-depth understanding of motivation and related theories. To this end, this section explores the use of motivational theories within the field in greater detail, paying particular attention to the most often used theories and drawing upon more recent efforts to unify the understanding of the role motivation plays in gamification.

Motivation theories can be classified according to two types of categories: natural (i.e., needs, drives, desires) versus rational (i.e., self-identity, meaningfulness, and instrumentality) and content versus process (i.e., 'what' vs 'how'). Rational theories include incentive theories of intrinsic (internal) motivation and extrinsic (external) motivation. Content theories refer to 'what' serves as a motivator, e.g., Maslow's hierarchy of needs (1967) or self-determination theory by Ryan and Deci (2000a). Process theories refer to 'how' motivation takes place, as in the goal-setting theory of Locke and Latham (2002) and the expectancy theory of Vroom (1964).

Intrinsic and Extrinsic Motivation

The most cited motivational framework in the studies explored concerned the distinction between intrinsic and extrinsic motivation, with 57% (n = 19) of the 33 studies citing the dissimilarity of the two. Extrinsic motivation is dependent on external factors and outcomes. For example, offering virtual money or points within a gamified system acts as an incentive for participants to carry out some desirable action. The aim of the extrinsic motivator is to produce approach behaviour, whereas the prospect of losing external elements may lead to defensive avoidance behaviour. Thus, the external incentive (money, points) elicits a behaviour which the participants are likely to associates with rewarding or punishing consequences (Ryan and Deci 2000).

Intrinsic motivation, the self-interest and satisfaction that drives an individual's actions, is viewed as a critical component of optimum experiences (Gagne and Deci 2005), i.e., influencing a user to undertake an activity because it is enjoyable is superior to employing external motivation as a motivator. Although the majority of gamification studies adopt this stance, discussions on methods to operationalise increased intrinsic motivation are lacking. Janitzek (2012) suggests the use of competition as a method for increasing intrinsic motivation but offers no validation. Whilst some individuals may indeed be intrinsically motivated by competition, its effects are not always positive (Johnson et al. 2005). Foster et al. (2012) state that, whilst intrinsic motivation appeared adequate for the majority of student subjects participating in their study, a significant minority required continuous prompting and extrinsic feedback or they simply lost focus, thus suggesting that external incentives may be a critical component in motivating some participants.

Mekler et al. (2013) proposes 'meaning' as the driver of intrinsic motivation, suggesting that providing a meaningful framing in the gamification application causes participants to experience a stronger internal urge to complete tasks associated with it than otherwise. Other researchers (Deterding et al. 2011, Nicholson 2015) in gamification share this view and call for meaningful interactions focused on increasing the perceived value of the activity to be placed at the heart of the process. Despite this, researchers such as Mollick and Rothbard (2013) and Herger (2014) describe gamification as a method for engaging workers in tasks without intrinsically motivating them and identify it as an evolution in management's relationship with employees and in how individuals do their work. Gamification offers a way to 'compensate' players through gameplay, argue Mollick and Rothbard (2013), thus rendering efforts to make the work more

meaningful no longer necessary. Although this position lacks empirical support, it appears in the marketing of many commercial products, where gamification is heralded as the solution to lack of engagement.

Whether a given motivator is perceived as extrinsically or intrinsically motivating depends on the individual and on contextual factors (Deci et al., 1999), and an objective distinction between these motivators, as some studies have suggested (Gnauk et al. 2012), is not possible. For instance, Zineddine (2012) and Xu (2011) amongst others have identified the game element badges as an extrinsic motivation. However, attaining badges may be perceived as intrinsic to the player, as found in a study by Firth (2013) on Foursquare. So whilst a reward may be viewed as extrinsic, the intrinsic desire for it may drive the participants' action within the activity.

Self-Determination

In exploring the motivational theories cited by gamification researchers, self-determination theory (SDT) was the most frequently cited. Developed by Richard M. Ryan and Edward L. Deci (1991, 1995, 2000a, 2000b), SDT was built on organismic theories which emphasise the person-environment dialectic and identify the three motivational factors within an individual's psychological needs as follows: autonomy, the need to be in control of one's life and act with self-volition; competence, the need to develop mastery; and relatedness, the need to empathise with, associate with, and relate to other people.

Gamification researchers interpret SDT inconsistently, particularly SDT's autonomy aspect. Whereas some view autonomy within gamification as simply giving the subject more designer-specified choices (Janitzek 2012, O'Donovan 2012), this interpretation within gamification may be at odds with the core concept of autonomy, where self-determined behaviour occurs when one's interests, preferences, and desires guide the decision-making process with respect to engaging in a particular activity. Heteronomy is performing activities at the direction of an external regulator, such as the designers of a gamification system. Whilst autonomy leads to feelings of control, imposing a heteronomic reward system on an individual may cause the individual to fail to perceive they are acting with autonomy (Sicart 2015).

Participant perception of competence or mastery appears consistently across the studies, perhaps as the concept of improving abilities is relatively straightforward and a general basis of many games. Gamification elements act as feedback to inform the user as to

status or score and thus act as a means to encourage improvement of abilities (Osheim 2013).

Social comparison theory, a part of cognitive dissonance theory, states that individuals seek to determine their social and personal standing based on how they compare against others whom they perceive as similar to themselves (Festinger 1954 as quoted in Vassileva 2012). Individuals then seek to enhance their abilities relative to their perception of standing amongst their peers, with maximum effort expended by an individual perceiving his performance as relatively lower than those viewed as inferior with respect to the specific ability (Suls et al. 2002).

Flow Theory

Flow has been demonstrated to increase performance in a wide range of activities (Csikszentmihalyi 1975, 1991; Demerouti 2006). Pates et al. (2003) describe flow as a state where skill and difficulty involved in achieving a task are in balance, i.e., a task perceived cognitively and affectively as neither too difficult, leading to anxiety, nor too easy, leading to boredom. Game-flow theory, developed by Sweetser and Wyeth (2005), provides a framework for quantifying play and enjoyment in games based on flow theory. Clear goals, control over one's activities, and immediate feedback are identified as antecedents to achieving flow in game and online platforms (Chen 2006, Sweetser and Wyeth 2005). A range of studies, including Eickhoff (2012), Korn (2012), and Ong (2013), utilise flow and game-flow theories to describe the principles behind their gamification designs, focusing in particular on appropriate challenges where the activity is neither too difficult nor too easy for the participant. Given the diversity of ability and interests of users of a system, the system would need to adjust the level of challenges in order to meet the user's optimum ability level. However none of the studies examined discuss or describe techniques to personalise the experience to different individuals, such that they experience 'flows.' Thus, whilst flow theory is cited by a range of researchers, since operationalized methodologies based on its assumptions are lacking, the theory is at best underutilised and at worst cited without merit.

2.1.2.6 Game Elements

The key components of most games are interaction, goals, rules, and rewards (Salen and Zimmerman 2004). Gamification uses similar components, in particular goals and rewards, as the principal methods of communicating the activities to be completed by the participant (Deterding et al. 2011, Seaborn and Fels 2015). In the gamification

studies examined, 72% (n=40) explicitly reported the game mechanisms utilised by the researchers.

Table 2.2 below provides a short description of each of the game elements together with their observed prevalence within the study sample. These game elements are also referred to as "motivational affordance mechanisms of gamification" by some researchers (Deterding 2011, Hamari et al. 2014, Nicholson 2012). Motivational affordance is derived from affordance theory (Gibson 1977), which describes the relationship between an entity that affords an opportunity for the subject to perform some action and the subject.

The overall findings are in line with similar studies in identifying the most common game elements used in gamification, i.e., points, badges, and leaderboards (Hamari et al. 2014, Nicholson 2013). In addition, role play, storyline, tangible rewards such as prizes, progressions, and challenges are also used to a lesser extent as noted by others (Nicholson 2013, Werbach and Hunter 2012).

Table 2.2: Prevalence of game elements within empirical studies on gamification.

Game Elements	Brief Description	Alternative Terms	Sample (n=40)*
Points	A unit for measuring or counting action or activity.	Experience points, karma points, social points, redeemable points, skill points, score	68% (n=27)
Badges	Visual icon denoting achievement.	Achievement, trophy	38% (n=15)
Level / Status	Increasing stages usually denoting overall progress. Can be numeric or textual.	Stage, title, rank, progress	35% (n=14)
Goals	Stated objectives or the aim or desired result of activity.	Objectives, challenges, quests	28% (n=11)
Leaderboards	Display of name of participants and associated scores.	Scoreboard, ranking	23% (n=9)
External Rewards	Physical or tangible desirable items.	Prizes, gifts, incentives	13% (n=5)
Role play / Story	The narrative premise of the activity.	Narrative, character	10% (n=4)

Game Elements as Feedback and Rewards

Incorporating game elements into an activity does not guarantee that the system's users will perceive it as having greater utility, being easier to perform, or being more fun to use. Fundamentally, gamification systems are additional layers added onto existing systems. Interactivity, which provides two-way communication between a subject and a system or other person, requires a system that provides feedback in response to the individual's actions.

Feedback is a cause-and-effect process and is commonly divided into two types: positive feedback and negative feedback. Although the term was coined in 1909 by Nobel laureate Karl Ferdinand Braun to describe coupling between components of electronic circuits, the terms is now widely used across many domains. In gamification, the term feedback is used interchangeably with rewards and is provided through such game elements as points and playing levels (Surugiu 2014, Codish and Ravid 2014).

In psychology, the effects of feedback can be traced back to behaviourist theories, as noted in the law of effect (Thorndike 1914) and operant conditioning (Skinner and Holland 1961). The law of effect states that positive feedback provides positive reinforcement of behaviour whilst negative feedback acts as punishment or negative reinforcement and that both improve performance. As Kluger and DeNisi (1996) noted, considerable scientific debate has occurred as to the degree to which rewards and punishments are effective in guiding behaviour, arguing that whilst the law of effect has considerable empirical evidence, it also has the advantage of parsimony and is too broad to explain the associated empirical complexities (Kluger and DeNisi 1996).

Reinforcement mechanisms such as points used in gamification are found in design frameworks for *persuasive technology*, the use of digital technologies to modify people's attitudes or behaviour (Fogg 2002). Various authors (Wu 2012, Zichermann and Cunningham 2011) have adopted these as a principle lens for viewing and designing gamification as a system of optimal stimuli and reinforcement (Deterding 2015). Reward and punishment as reinforcement mechanisms have a long tradition in behavioural psychology, from the "Skinner Box" (Skinner 1953), wherein animals placed in a box are rewarded or punished depending on either a time or some other specific response, such as pressing a lever for food. However, equating such reward-based elements as points and leaderboards to a reinforcement framework has drawn strong criticism of gamification as a modern version of a Skinner Box (Bogost 2011b) which make extensive use of both positive reinforcement, as in gaining points or

position in leaderboards, and of negative reinforcement or punishment, as in losing points or ability. For example, in Farmville (Chang 2012) if crops are not harvested within a certain time interval, they die. Within gamified application, designers have tended to make minimal use of aversive contingencies (situations that players wish to avoid), fearing that these may lead to disengagement (Linehan et al. 2015). This rather one-sided use of reinforcement may help explain why some gamified systems fail to engage (Foster et al. 2012). Critically, in order to apply reinforcement techniques effectively, the nuanced difference between a reward, a stimulus given to a participant on the *assumption* that it increases the chances of a behaviour being repeated, and its specific reinforcer, a stimulus that *has been observed* to increase the intended behaviour, must be understood and explored. Researchers have argued that, in order for reinforcement mechanisms such as points and leaderboards to be effectively applied in gamification, careful observations and measurement of behaviour as to specific consequences of use of reinforcers or punishers must be captured and analysed rather than simply assumed and that this should constitute an integral part of the development process (Linehan et al. 2015).

Feedback is a vital component of any learning environment and a critical feature of interactive games. Feedback supports self-efficacy and self-regulation and provides direction and guidance for performance (McNamara et al. 2009). McNamara et al. (2009) highlight three forms of feedback within digitally enhanced learning systems: information-feedback, where a specific feedback provides insight as to quality or accuracy of a user's actions; consequence-feedback, where the system reacts to a user's action as, for example, when a wrong action by the player leads to that player losing a life; and point-feedback, where points can be cumulated over multiple performances through progress boards or levels to indicate directly or indirectly the accuracy or quality of a user's actions.

Some of the studies that provided evidence for improvement in performance reported substantial improvements, as, for example, the experiment by Berengueres et al. (2013) on rate of student recycling, which measured a 300% improvement in the rate of recycling by providing a simple feedback mechanism, thereby demonstrating that a gamification system does not have to be complex in order to affect behaviour. Indeed, the literature on changing behaviour suggests that small changes can result in profound differences in outcome (Thaler and Sunstein 2008).

Feedback or rewards can also have negative consequences, as, for instance, in the study of Hamari and Koivisto (2013) where badges are provided on a continuous basis without denoting any real achievement. As demonstrated by Lawrence et al. (1962), faced with a monotonous outcome for an activity, "the organism discontinues the activity." Hamari and Koivisto (2013) observed that the group provided continuously with badges did not exhibit a statistically significant improvement in performance, apparently contradicting Firth's (2013) observations that awarding of badges increases performance. However, in Firth's (2013) study, receipt of badges were generally regarded as significant achievements by the players, requiring the completion of a variety of activities, with many badges being deemed as rare or difficult to earn. In contrast, the experiment of Hamari and Koivisto (2013) rewarded participants automatically and continuously with little variety, hence causing them to be perceived as less valuable or significant.

In studies where performance was seen to decrease, insufficient attention to the user-experience design and lack of deep expertise in interaction design on the part of the software developers appears to have been a factor in at least some of the experiments conducted; comparing the interfaces presented to users in these study experiments with established design practices (Crumlish and Malone, 2009) supports this hypothesis.

Points in Gamification

In the gamification studies explored, 73% (n=40) utilised points as a feedback or reward mechanism. In one of the largest empirical studies, polling 3,500 participants, Thom et al. (2012) found that removal of a point-based incentive system significantly reduced participants' overall contributions ($F(1,1717)=38.24$, $p<.0001$), thus suggesting a strong link between gamification and the motivational affordance offered by points, particularly where points are used as a method of providing utility.

However, the tendency of gamification systems to rely on point accumulation, i.e., numeric units earned as a result of performing an activity (Seaborn and Fels 2015), does not always produce a positive behavioural outcome (Deterding 2014), particularly where points are devoid of meaning in so far as they fail to support the user's personal goals or genuine interests (Mekler et al. 2013). Indeed, researchers have used the term 'pointsification' in a derogatory sense (Kotini and Tzelepi 2015, Rughinis 2013) to identify the practice of gamifying a system simply by adding points rather than developing the rules or structure that would make a real game interesting.

Awarding points for the performance of 'right' behaviours as deemed by the system designers is a subject of serious derision by gamification researchers; for instance, Nicholson (2015) compares it to treats given to dogs to get them to behave. Indeed the motivation model may be viewed as resembling Skinner's reinforcement theory (Skinner 2014), where extrinsic rewards reinforce specific behaviours. The pattern of awarding points (Lillienfeld et al. 2009) can elicit different performance levels. For instance, variable reward schedules, where rewards are given at non-fixed time intervals, appears to result in a higher level of performance in comparison to a fixed reward schedule, wherein the participant knows specifically when the next reward will come (Lillienfeld et al. 2009). All types of gambling operate on a partial or variable reinforcement schedule, where the player is periodically surprised by the outcomes (Sharpe and Tarrier 1993). The notion of surprise is also, according to Malone (1981), deemed a critical component of engaging the curiosity of learners, and recent research has demonstrated that the element of surprise enhances infants' learning and exploration (Stahl and Feigenson 2015). Despite the mounting evidence that employing variable rewards has the potential to increase performance, given their semantic proximity to notions of rewards in a gambling context and the possible perception of unfairness (i.e., when two students carry out the same activity at the same level of competence but receive differing reward amounts), their use remains limited. None of the studies reviewed appeared to have used variable rewards.

Mekler et al. (2013) found that points used as an extrinsic motivator appeared to increase the quantity of tags produced but had no impact on the quality of the tags generated by the participants within a gamified system. However, by associating a specific meaning or narrative to the effort to accumulate points, for example by framing the activity in terms of supporting a non-profit organisation, the most amounts of time and the highest quality performances resulted. Ong (2013) found a similar effect when comparing and combining the use of points and narrative, revealing that employing points alone increases quantitative measures of task performance while narrative increases intrinsic motivation and quality of output.

A design employed in a study by Burkey et al. (2013) increased cooperation among team members and improved student attitudes within a laboratory course, although no significant improvements in student grades were achieved. The design incorporated points of two types, recognising reputation and experience, depending on whether quests or other achievements were completed. The sum of the team members' points

was then normalised by the number of students within each team, thus ensuring that an individual's success ultimately rested on the totality of all team members' contributions, and so provided a reference and framework for increased cooperation.

Thus, whilst points as a feedback mechanism enable users to gauge their performance, the context and meaning of the points themselves change the perception of their efficacy and regulatory properties. Points can be used in a variety of forms—as progress indicators such as progression bars; as a method of achieving rankings and so leading to or triggering competition (Liu et al. 2011); as a means of measuring reputation (Prause 2013) such as Karma points, which may lead to cooperation or competition; or simply as a way to encourage self-mastery without any social comparison (Richter et al. 2015).

Goals in Gamification

Of the studies examined, 28% (n=11) used goals explicitly as a gamification element. Goal-setting theory as an approach to increasing performance first appeared in the 1930s (Latham and Pinder 2005). Locke (1968) laid the foundation for goal-setting theory by establishing a positive relationship between identified goals and performance. Locke and Latham (1990) demonstrated that goals can increase performance by motivating participants to exert greater effort, encouraging them to persist, and guiding their actions toward specific outcomes. Furthermore, goals where performance outcomes are given feedback publicly (within the group) have been shown to strengthen group members' commitment by stimulating upward social comparison (Hollenbeck et al. 1989, Jung et al. 2010, Mowday et al. 1979). Meyer (2003) further developed a goal-setting framework that maximised the effect of goal setting through use of goals that were specific, measurable, achievable, realistic, and time targeted.

The Hamari (2013) study examined the impact of goal-setting on quantity and quality of users' actions within a gamified system and observed no relationship between goals set by the system and output. Lack of utility afforded by the goals presented within the system and the introduction of goals long after the launch of the service, thereby making them unexpected and unfamiliar to users, may account for this outcome. Mollick (2013) cites the heteronomous or 'mandatory' nature of the goals within a gamified system as a potential negative effect possibly leading to loss of motivation.

In a study of virtual achievement by Denny (2013), goal-setting resulted in a highly significant positive effect on the quantity of participants' contributions without a corresponding reduction in their quality. Specifically, use of goals resulted in a 22%

increase in number of questions answered (Mann-Whitney U=116386.5, Z=3.45, WithGoals=516, WithoutGoals=515, p<.001). In examining the effects of goal-setting across the studies, it is the utility of goals that drive motivation.

Leaderboards in Gamification

A leaderboard is a game mechanism consisting of a visual display of users or groups ranked according to their score or accomplishments. Gamification researchers have described leaderboards as supporting the need for status, recognition, belonging, and prestige in order to strengthen competence and mastery (Dominguez et al. 2012, Mollick et al. 2013, O'Donovan 2012). The general nature of leaderboards provides a competitive framework where users compete to obtain higher ranking based on their efforts. However as highlighted by Nicholson (2013), while the presence of leaderboards could encourage some, it could discourage others; those ranked high on the leaderboard may enjoy a sense of superiority which prolongs their engagement whilst those at the bottom may be demotivated and give up as they perceive little chance of catching up.

Social comparison theory has been used to explain the motivational effects of leaderboards (Heinzen et al. 2015, Vassileva 2012). In the context of social comparison theory, upward comparison (comparing oneself with those ranked higher than oneself on the leaderboard) may offer hope or inspiration, i.e., a reason for increasing one's effort for higher achievement (Buunk et al. 2005, Collins 1996), whilst downward comparison (comparing oneself with those ranked lower than oneself on the leaderboard) may provide a sense of pride or achievement. However, researchers have argued that social comparison may elicit significant negative consequences, whereas upward comparison may lead to feelings of jealousy or injustice as one's value is determined to be lower than that of another, contradicting the individual's perception of himself as in fact worth more. Downward comparison may provoke negative feelings or resentment, as one individual perceives himself as working much harder than others, who are free riding, particularly if the individual views their potential future success as predicated on the actions of those others (Buunk et al. 2001, Buunk et al. 2005). Other studies concerning effects of leaderboards in gamification on academic achievement support social comparison processes, empirically demonstrating that leaderboards can produce unexpected effects which negatively impact performance (Christy and Fox 2014, Hanus and Fox 2015).

Thus, the presence of leaderboards, as with other game elements in gamification, does not guarantee increased motivation or performance. Instead, their impacts may, to a greater extent, be linked to their perceived meaning by the participant driven by the way in which they are framed and communicated.

2.1.2.7 Unintended Outcomes

Gamification of astronaut training, the topic of an empirical study by Cornelissen et al. (2012), found that not only did gamification reduce user motivation but it also had a detrimental effect on training outcomes. Review of this study indicated that the cause of the negative outcome was lack of appreciation for the intended audience in the system's design. Some participants, particularly those who were not used to such game-like features as achievements, found the application too child-like and were demotivated by the sense of micro-achievement associated with this type of training. These users, many of whom appear to have been older astronauts and trainers, argued that learning to use the astronautical equipment *"should be"* enough motivation for prospective students and that the inclusion of gamification detracts from the seriousness of the learning activity.

Dumitrache et al. (2013) found that, rather than improving outcomes directly, the inclusion of gamification shifted the focus of the participants from creating new tags to evaluating tags that others had suggested. Enabling participants to access and comment on each other's contributions resulted in fewer tags being generated and a reduction in variety as participants sought to evaluate one another's contributions rather than contribute a fresh tag. Unintended consequences are an important aspect of gamification, as additionally noted by Xu (2011) and Landers et al. (2011), possibly including increased unethical behaviour or reduced motivation (Locke and Latham 2009, Zhao et al. 2010).

Kani et al. (2013) studied the effects of gamification on CAPTCHA, which asks a user to verify that they are human rather than machine. The impact was an increase of 28% (n=10) in user satisfaction in the gamified system contrasted with the non-gamified one. This increase in satisfaction did not, however, result in increased performance since the gamified system took 30 seconds per action compared with the 5 seconds per action for the non-gamified version. Participants attributed the increased satisfaction to the increased entertainment value of the activity. Therefore, in this instance, whilst gamification improved task satisfaction, it was detrimental to overall performance.

2.1.3 Summary of Gamification Review

Exploring the sample of empirical studies on gamification revealed that most lacked a robust methodological approach; this lack manifested as not including comparison groups, overreliance on a singular measurement, short treatment, and absence of validated measures, observations echoed by Hamari et al. (2014) and Heinzen et al. (2015).

Whilst generalisation given the diversity of the studies is difficult, in summary, of studies that demonstrated an improvement in outputs, the emerging theme appears to be enabling participants to better structure and verify their activity whilst receiving feedback within a social setting.

Leading researchers in gamification (E. T. Chen 2015, Hamari 2013, Nicholson 2015) have identified the importance of building a community of committed individuals aligned with the gamification's goals as one factor critical to its success. They note that enabling exposure by participants to the attitudes of others and allowing participants to give and receive feedback through the gamification positively influences their motivation. As explored in goal-setting theory (Locke and Latham 1990), commitment brings forth a willingness to promote and recommend the service and thereby increase retention and acquisition of new users. In practice, Hamari (2013) suggests imbuing gamification with elements that increase social interaction among the cohort, in particular, by incorporating actions that enhance social influence and the perception of reciprocal benefits.

In summary, studies which displayed the greatest increase in performance tended to share an appreciation for good social design patterns where the needs and goals of the users, as part of the community of users, form the heart of the design. Developing a meaningful game-based experience that connects the underlying non-game settings and activities with the user in a pragmatic manner forms the focus of Nicholson's (2012) theoretical framework for meaningful gamification.

2.2 Interdependence, Cooperation, and Competition

Research literature on cooperation and competition has a long history beginning in the late 1800s. Triplett (1897, as quoted in Johnson and Johnson 1974) demonstrated that performance of both children and adults improves when they are placed in groups rather than isolated from others.

In the early 1900s, Kurt Koffka's theory that groups were dynamic wholes in which interdependence between members could impact their effectiveness gave rise to field theory. Between the 1920s and 40s, Kurt Lewin (1935, 1948) refined Koffka's notions of interdependence further with the principle of contemporaneity. Where social behaviour is inherently contextual, an individual's understanding of the environment he perceives himself as contending within determines his actions as his behaviour unfolds. Since life is dynamic, an individual's perceptions of situations change according to perceptions of desired goals and the environment (Johnson and Johnson 2005).

Lewin proposed that the essence of a group is interdependence among members and the perception of common goals in conjunction with collective motivations. In interdependence, group members share common goals, resulting in a dynamic interaction such that a change in state by any one member of the group changes the states of the other members. Moreover, an intrinsic state of tension exists among group members and motivates collective movement toward the accomplishment of the desired common goals. In still further research, Lewin contended that the drive for goal accomplishment motivates cooperative and competitive behaviour (Tjosvold and Johnson 2000). Social interdependence exists when individuals have overlapping goals, and each individual's outcomes are affected by the actions of the others (Deutsch 1949a, 1962; Johnson and Johnson 1989; Johnson 2003).

2.2.1 Theory of Cooperation and Competition

Although the terms *cooperative* and *competitive* are sometimes used to denote individual dispositions or behaviours, the concepts in their intended frame of reference denote perceptions of a given situation by those taking part, including the consequences in relation to others that their behaviour would entail. Extending Lewin's reasoning about social interdependence, Deutsch formulated a theory in which cooperation and competition are two opposing types of social interdependence forming a continuum (Deutsch 1949a, 1962), where cooperation can be thought of as *positive interdependence* and competition as *negative interdependence*. The absence of social interdependence and dependence results in *individualistic efforts*.

In the eleven decades since Triplett (1898) conducted his studies on competitive performance, more than 1,200 studies have explored the relative merits of cooperative, competitive, and individualistic efforts within small groups, representing one of the largest bodies of research within the field of psychology. Building on the works of

Deutsch, in the early 1970s, researchers David Johnson and Roger Johnson began theorizing about and researching cooperative, competitive, and individualistic efforts. Over the last 40 years, Johnson and Johnson have published over 100 research studies on this topic, particularly as it relates to learning and education (Johnson and Johnson 2009).

In formulating the theory of competition and cooperation, Deutsch (1949a, 1962) made a number of assumptions, one of which was that goals are defined as singular instances, objectively categorisable into one of the three categories—cooperation, competition, and individual. However, in the complexity of the real world, individuals often have multiple goals, and so cooperative, competitive, and individualistic goals may be present simultaneously in the same situation. In Deutsch's theory, groups were assumed to be small in size, not necessarily a valid assumption in the real world. In addition, it assumed that all participants possessed equal power, whereas the distribution of power, almost always unequally distributed, has been shown to have considerable effect on motivation (McClelland 1987) and productivity (Tjosvold 1990). The theory assumed individuals to act in their best self-interest rather than imitating others in herd-like behaviour, as has been shown to be the case in reality and thus to impact actions (Earls 2009). Deutsch further assumed that a situation was independent of past events, rarely the case in real life (Johnson 2005).

2.2.2 Positive Interdependence (Cooperation)

Positive interdependence exists where the structure of interactions between individuals is positively related, such as when participants within a group can only succeed if *all* group members succeed, i.e., effectively a collaborative situation where "all members will sink or swim together" or win-win as suggested by Johnson and Johnson (1998).

Positive interdependence can occur through outcome interdependence, means interdependence, or boundary interdependence (Johnson and Johnson 2005). Outcome interdependence consists of goals interdependence and rewards interdependence. Means interdependence is said to exist if one or more dimensions of a resource, role, or task are interdependent within the team.

Positive goal interdependence exists when individuals perceive the attainment of their own goal as predicated on all other individuals with whom they are cooperating. Johnson et al. (1991) reported positive goal interdependence as resulting in increased

performance through higher achievement and greater productivity in contrast to resource interdependence (described below).

Positive reward interdependence exists when the incentive structure is inherently cooperative, such that group members receive a collective reward or each receives a reward equal to that of other members for successfully completing a joint activity. Johnson and Johnson (1989) identify positive reward interdependence as having the highest positive impact on task-related efforts where all individuals contribute to attaining the collective reward. Furthermore, performance level increases when both goal and reward interdependence are present (Lew et al. 1986, Mesch et al. 1988).

Resource interdependence exists when no one individual within the collaborative team controls all resources, information, or materials needed to complete the task. Thus, resource interdependence requires members to combine resources in order to complete the collective objective.

Oritz et al. (1996) demonstrated that resource interdependence not combined with other positive interdependences can negatively impact performance in contrast to individualistic effort. When individuals share resources without having common goals or rewards, their focus shifts more toward obtaining the resources of others and less toward sharing their own, thereby reducing overall performance level (Johnson et al. 1990).

Role interdependence exists where individuals' capabilities and responsibilities are complementary or interconnected with the roles of others as, for example, in an operating theatre where surgeons, consultants, anaesthetists, and nurses work together to complete an operation.

In summary, positive interdependence encapsulates situations in which the group shares a common reward and strives for a mutual benefit, often within a long-term time frame. Identity is shared, causation mutual and affiliation motives are enhanced where each participant views himself as critical in the performance and outcome of the group. Thus, the sense of mutual responsibility and mutual obligation to support and assist is characterised by positive inducibility and mutual investment in other team members (Johnson and Johnson 1989).

Positive interdependence has been shown to increase participant motivation to increase performance level through both increased effort and increased cognitive reasoning strategies (Gabbert et al. 1986). Furthermore the stronger the interdependence through

goals, rewards, and means, the greater the sense of group cohesiveness and unity felt by the entire team (Lickel et al. 2000, Wellbourne 1999). Thus, awareness that one's performance affects the group's success increases the sense of responsibility experienced in exerting effort.

2.2.2.1 Social Loafing in Cooperation

Cooperation can also provide opportunities for social loafing. "Social loafing is the tendency for individuals to expend less effort when working collectively than when working individually" (Karau and Williams 1993). Karau and Williams (1993) developed the collective effort model through meta-analysis of 78 studies on social loafing and by drawing on social impact theory (Latane 1981) and social identity theory, which proposes that individuals' sense of identity is based on the social groups they associate with (Abrams and Hogg 1990, Tajfel and Turner 1986) and group-level versions of social comparison theory.

Karau and Williams (1993) determined that individuals are less likely to engage in social loafing when their output is evaluated collectively, when working on tasks that appear meaningful, when group-level comparison is available, and when their input to the collective outcome is dependent on those of others in the group. In short, the propensity to socially loaf is higher in cooperative structures where an individual's efforts are not easily distinguished from that of others.

2.2.3 Negative Interdependence (Competition)

Negative interdependence exists where the structure of interactions between individuals are negatively related so that individuals perceive their success as dependent on others' failures. Effectively this represents a competitive situation where outcomes are win-lose, where one succeeds only at the expense of another.

Negative outcome interdependence exists when either goals and/or rewards are mutually exclusive, i.e., more success for one means less success for others. For example in a race, the entity that achieves first place denies all others the ability to also place first (Johnson and Johnson 1989).

Negative means interdependence exists when the resources, roles, or information required to achieve a certain outcome are mutually exclusive, such that actions of one individual directly prevent others from accomplishing the same action at the same time. Competition may exist with or without means interdependence, and the form of

negative means interdependence varies with circumstance but, in general terms, can be viewed as aimed at opposing, contesting, overcoming, or otherwise eliminating or working against others.

Thus, negative interdependence incorporates the recognition that to 'win' means working against others, as one's gain is another's loss and helping another individual could lessen one's own chances of winning. The time frame in negative interdependence is often short term and identity is negatively cathect, especially in win-lose situations. A lack of inducibility, i.e., readiness to accept another's influence, means competitors may attempt to obstruct one another (Deutsch, 2006).

Working and learning in teams comes naturally to humans, as social creatures evolved to survive based on mutual cooperation. Even competition, especially within learning or work environments, can be viewed as a type of cooperative activity, in so far as participants cooperate with respect to essential rules, context, and duration of activities. Underlying agreement on the rules and boundaries of competition can aid in enhancing outcomes, where winning and losing against other participants is not the primary objective, as for example when friends compete in a maths lesson. In these situations, competition plays a healthy, constructive role as the nature of the shared experience provides the potential for excitement without being detrimental to the overall performance of the group.

The view of competition as constructive is not universal, as critiques of competition point to its causing anxiety, aggression, self-doubt, and lack of effective communication (Kohn 1986). However, humans have a tendency to compete in activities such as games and sports, and the continued popularity of such activities would argue against their innate destructiveness. The right type of competition can not only increase motivation and commitment in the face of challenging tasks but can also increase the desire to maintain membership in a group.

However, for competition to be effective, certain structural elements that support participants must be addressed and managed; these include appropriateness of tasks, measures to address anxiety related to losing, and a means to address inappropriate behaviour, all of which are intended to ensure a contest appropriate for the skills and abilities of the competitors (Johnson and Johnson 1989). Possible negative effects of competition occur when participants lack the appropriate competitive skills for the contest, handle winning or losing inappropriately, fail to play fairly, over generalise

outcomes, or lack perspective as to the true purpose of the competition, i.e., using it to improve ability by increasing skills and knowledge.

2.2.3.1 Complementary Theoretical Approaches to Competition

A relatively recent contribution to competition theory is Ryckman et al. (1990, 1996, 1997) in which a theory of personal competitive orientation supporting two distinctive types of competition is proposed. In the first, *hypercompetition,* achievement of personal outcomes is carried out with little concern for the means used and possible harm to others incurred in doing so. On the other hand, *personal development competition* views competition as facilitating personal growth, where individuals compete without demeaning, demonizing, or otherwise acting in a detrimental manner with respect to their competitors. Personal development competition could be viewed as having two goals: to win and to use competition as an opportunity for personal growth such that others are not harmed.

Tjosvold et al. (2006) adopt an approach similar to that of Ryckman et al. in defining two types of competition. The authors term their two types *constructive competition* and *destructive competition.* Constructive competition has characteristics similar to those of personal development competition, that is, the desire to compete without depriving others or acting with intent to stop others from succeeding. Indeed, a series of studies (Tjosvold et al. 1992, 1998, 1989, 2006) has shown constructive competition as facilitating emotional enjoyment, task effectiveness, social support, and ability to continue working collaboratively with competitors in the future.

Ryckman et al. (1996) employ similar concepts with respect to team-based competition. They term the two types of competition as *team hypercompetition* and *team development competition*, in which individuals compete within a team without resorting to unfair tactics, hostility, or high levels of anxiety, and a higher level common goal or achievement drives participants to maximise their own outcomes constructively whilst at the same time improving the overall team performance (Tjosvold et al. 2003). Thus, an individual engaged in team development competition is primarily focused on collective growth and team mastery leading to learning and self-improvement (Collier et al., 2010). In team hypercompetition, on the other hand, individual competition could become a source of conflict leading to an intra-team zero-sum game negatively impacting performance and outcomes (Johnson et al. 1981).

Additional theoretical approaches to competition and cooperation outside social interdependence theory include social exchange theory, proposed by Kelly and Thaibaut (1959, 1978), which defines competition as a means for individuals to maximise their rewards whilst minimising their costs with respect to others. Furthermore, social learning theory of group agency, a subset of social cognitive theory developed by Bandura (1999, 2001), distinguishes between three modes of agency: personal, proxy, and collective, where the duality between personal and social foci of causation gives rise to bidirectional influence between social structure and personal agency.

2.2.4 Social Interdependence and Performance

Deutsch's proposed theory of competition and cooperation has provided the conceptual framework to organise and analyse hundreds of research studies in diverse fields including education and business (Johnson and Johnson, 2005). The last five decades have seen a number of systematic reviews of the research and evidence on the field published. Johnson and Johnson (2005) identified 754 studies published between 1903 and 2005 which explored the relative merits of cooperative, competitive, and individualistic efforts and the conditions under which use of each is appropriate.

Johnson et al. (1981) reviewed 122 studies comparing the relative effectiveness of goal structures based on cooperation, competition, and individualistic efforts in promoting achievement and productivity involving such lower level cognitive tasks as motor skills, decoding, and recall of factual information. The authors presented three critical findings. Firstly, cooperation was deemed considerably more effective than interpersonal competition and individualistic situations. Secondly cooperation in intergroup competition performed better than interpersonal competition and individualistic efforts. Thirdly, no significant differences in results were observed between interpersonal competition and individualistic effort situations.

Building on the Johnson et al. (1981) analysis, Qin et al. (1995) reviewed 63 studies comparing the impact of cooperative, competitive, and individualistic efforts on problem solving across a range of situations involving higher level cognition, both linguistic and nonlinguistic activities, and well-defined and ill-defined operations. The overarching finding was that individuals within cooperating teams outperformed individuals competing with each other on all types of problem solving in 55 of the studies, with 8 studies reporting that competition outperformed cooperation. On

average, the meta-analysis showed that cooperators outperform hyper-competitors by 1.5 standard deviations.

More recently, Hilk (2013) carried out an extensive meta-analysis of 231 experimental studies involving 37,422 college and graduate-level students to investigate the effects of competition, cooperation, and individualistic goals on achievement and peer relationships. The study showed that cooperative conditions positively promote achievements and peer relationships to so significant a level in comparison with hyper-competitive or individualistic settings that not only an expert but even a casual observer would notice the difference. With respect to peer relationships, this finding follows the theory of social interdependence in that most evidence supports use of positive relationships in cooperative conditions. However, cooperative versus competitive conditions produced a greater variance between effect sizes for both achievement and peer relationships whereas comparing individualistic conditions with competitive or cooperative conditions resulted in lesser effect sizes.

2.2.5 Game Theory Perspective of Competition and Cooperation

Game theory plays an important role in many domains of economic theory, as it is a mathematical representation of equilibrium and transfer of assets between actors in markets, contracts, and political economies.

Employing game theory in modelling competition, cooperation, and non-cooperation games allows exploration and testing of a variety of zero-sum and non-zero-sum situations and their outcomes. Game theoretic zero-sum competition bears the closest similarity to negative interdependence, where both outcome and means are interdependent, as conceptualised by Johnson and Johnson (1989); to hypercompetition as conceptualised by Ryckman et al. (1997); and to destructive competition as conceptualised by Tjosvold et al. (2006). Constructive competition or personal development competition can be viewed as non-zero-sum competition in that individual participants can succeed without detriment to others.

The exacting nature of representing competition or cooperation within classic game theory assumes sophisticated and highly logical actors capable of maximising profits and utility in a given situation. Whilst game theory constructs are often tested through simulations and provide some level of predictive ability, particularly when large populations are being considered, psychological testing and experimental evidence with

smaller groups has led to the development of *behavioural game theory*, where learning, social preferences, and other behavioural biases lead to outcomes that seem counter to the notions of a Nash equilibrium or utility maximisation (Eyster and Rabin 2005).

Since studies in classic game theory focus on bargaining among rational players within binding agreements rather than on individuals exhibiting wildly varying personality traits and subject to mood changes and irrational behaviour, the scope for investigating behavioural game theory in the context of this study makes it too limited to be considered. Thus, although behavioural game theory and *evolutionary game theory*, developed for study of cooperation and competition based on behavioural economics, are relevant to this study, given their focus on large populations, their scope for application in small groups is limited (Charlesworth, 1996).

Thus, in order to better define what is a meant by competition or cooperation, developing a unified approach was necessary.

2.2.6 Unified Framework for Competition and Cooperation

Through combining social interdependence theory (Deutsch 2006), extensions on competition developed by Ryckman et al. (1990, 1996) and Tjosvold et al. (2006), zero-sum and non-zero-sum game theory (Axelrod 2006, Von Neumann and Morgenstern, 2007), and social exchange theory (Kelly and Thibut 1978), construction of a unified model of competition and cooperation is possible. Moreover, such a model can distinguish interdependence between means (i.e., abilities, information, resources, and skills) and interdependence between outcomes (i.e., rewards, recognition, and goal achievement).

This unified approach can, at least partially, explain the apparent disparity between the evidence found by different researchers in support of or in opposition to competition, as it offers a more nuanced approach in defining the specific social interdependence within a given study. This approach can also, to some extent, explain the findings that competition facilitates higher motivation and increased performance as suggested by Abuhamed and Csikszentmihalyi (2009) and Tauer and Harackiewicz (2004), who appear to have been examining personal development competition, against negative effects on group cohesiveness and friendship (Tjosvold et al. 2003), where profound competition was being considered. Similar disparities in effect with respect to cooperation also remain; despite a vast collection of research on the effects of

cooperative learning, major disagreements on which structures lead to the best outcome gains for participants continue to exist (Johnson et al. 1981, Sharan 1990, Slavin 1983).

Table 2.3: A Unified Framework of Competition and Cooperation.

		Outcome (Rewards) Interdependence		
		Positive	**Neutral (no interdependence)**	**Negative (zero-sum)**
Means (Resources) Interdependence	**Positive**	[1] Coordinated interaction toward achieving mutual rewards. 'We work together sharing resources, skills; we win/lose together.' **Hypercooperation**	[2] Coordinated interaction toward independent outcomes. 'We work together sharing means to achieve outcomes that are independent.' **Personal Development Competition (PDC)**	[3] Coordinated interaction toward contrariant outcomes. 'We work together but compete on outcomes.' **Contrariant Cooperation (scarce rewards)**
	Neutral (no interdependence)	[4] Means independent interaction for shared outcomes. 'We work independently; we win/lose together.' **Outcome Cooperation**	[5] Entirely individualistic. 'I work independently, and I succeed independently.' **Individualistic** (may be interpreted as PDC)	[6] Means independent for competitive outcomes. 'We work independently, compete on outcomes.' **Outcome Competition**
	Negative (zero-sum)	[7] Oppositional interaction toward mutual outcome. 'We compete on means, we win/lose together.' **Contrariant Competition**	[8] Oppositional interaction toward independent outcomes. 'We compete on means but succeed individually.' **Means Competition**	[9] Oppositional interaction toward contrariant outcomes. 'We compete on means and outcomes.' **Hypercompetition**

[1] **Hypercooperation** is marked by complete sharing, i.e., positive interdependence between both means and outcomes exists (Johnson and Johnson 1989). This type of cooperation provides the greatest opportunity for optimal team performance, provided such contextual factors as social loafing (Karau and William 1993) are managed. Example: *Players in a football team. Individuals must work together, drawing on each other's strengths, and the team wins or loses as a single entity. In an educational setting, this would be akin to group marks awarded to all members of a class team for their combined effort without singling out individuals.*

[2] **Personal Development Competition.** Individuals do not impede others from achieving their outcomes (Ryckman et al. 1990) but may share means or be interdependent in some other way, i.e., rely upon or use each other's abilities. Example: *A typical class-room setting, where students may work together, sharing resources and information. However, the grades students receive are not dependent on the efforts of others class members, and so everyone could receive full marks.*

[3] **Contrariant Cooperation**. Individuals work together, sharing means, but compete with respect to outcome. This type of cooperation may have destructive effects as, whilst individuals need each other, their outcomes are negatively interdependent. Example: *Winning a Nobel Prize. Whilst many individuals may have contributed to the overall effort leading to the discovery, only one or two individuals receive the 'reward' of being awarded the Nobel Prize itself.*

[4] **Outcome Cooperation**. Individuals belong to a team but work independently despite sharing in outcomes. All team member perform the same actions, and these actions have no direct effect on the performance of other team members; however the overall outcome is considered to be the sum of individual team member's efforts. Example: *A chess team, where players of a team play players from other teams and the overall outcome is at the team level.*

[5] **Individualistic** situations are neither competitive nor collaborative. However, framing this situation in terms of personal development competition can be accomplished by simply providing the ability for individuals to compare their efforts with that of others. Example: *Students from different institutions, who have presumably had no previous interaction with one another, sit the same exam. As part of the framework in such a personal development competition, the students' grades would be ranked in a leaderboard.*

[6] **Outcome Competition**. Individuals work independently but compete on outcomes. Example: *Track and field sports such as the 100-meter dash. Whilst each runner relies on his or her own means, each competes with respect to finish position (i.e., outcome).*

[7] **Contrariant Competition**. Individuals within a team compete on resources and share an outcome. Example: *Within a class room where the group is allocated a single computer (i.e., the resource), only one or two individuals can use the resource at any one time but the 'reward' (i.e., grades) are awarded at a group level. This situation requires a high degree of coordination, as limited resources could lead to negative*

competitive behaviour between members of the group and thereby be detrimental to everyone's outcomes.

[8] **Means Competition**. Individuals compete only with respect to means or resources but not to outcomes. Competing on means seldom has no impact on outcomes, but one possible case is employees whose salaries are not performance related but who nevertheless have limited resources with which to perform their jobs. Example: *A single piece of machinery or meeting-room is available, so that if the machinery or meeting-room is taken by another employee, the individual's capacity to maximise his effort is reduced, but individual outcome, in this case salary, is not affected.*

[9] **Hypercompetition**. Individuals compete on both means and outcomes. In game theory, this is termed *strictly competitive* (Axelrod, 2006). This type of competition is considered destructive within social interdependence models (Tjosvold et al. 2006). Example: *Two-player chess matches, where players battle over means and outcomes.*

Whilst any one of these nine potential interdependent situations could occur in a social situation involving tasks and outcomes, real world situations may consist of an amalgamation of interdependencies and thus be more complex to categorise.

2.2.7 Summary of Cooperation and Competition

As has been seen, much of the research exploring the differences between competition and cooperation fails to account for the more nuanced interactions between positive, neutral, and negative interdependence of means and outcomes, thus leading to widely differing claims against and for either cooperation or competition. As demonstrated, there are many different types of competition and cooperation, and real-life situations are often of such complexity that more than one type is present thereby forming an amalgamation of different types. For instance, whilst students may compete for the teacher's time (i.e., a resource) under hypercompetition, they may cooperate by sharing information with one another, while at the same time also competing with respect to individual outcomes (i.e., personal development condition).

Added to this mix of possible interdependencies are the often neglected dimensions of personal preference and personality. As theorised by Bandura (1986, 1999) in social cognitive theory, the individual's personal agency, including intentionality, form part of the interplay among the individual's personality, the socially interdependent situation, and the wider context. Additionally as shown in studies of personality within the

context of team performance (Chen et al. 2011) and contextual performance (Borman and Motowidlo 1993), individuals differ with respect to beliefs and attitudes regarding the nature of their relationships, the context, and cooperation and competition themselves.

Therefore, further study of the effect of personality on preferences in social interdependence type may provide further insight into the differences in outcomes and performance measured by different researchers when examining the effects of competition and cooperation, this is the topic of the next section.

2.3 Personality

Personality is the combination of characteristics and qualities that form an individual's distinctive character. Personality traits refer to the enduring personal characteristics revealed in a particular pattern of behaviour displayed in a variety of situations (Allport 1937, Costa and McCrae 1992, Goldberg 1993, McCrae et al. 2000).

During the search for an ideal personality model for use in this thesis, a number of models were considered. A search through the games research domains revealed Bartle's four personality types as the most prevalent model (Bartle 1996). Here, players are categorised as Achievers, Explorers, Socialisers, or Killers. However despite being one of the most often-cited models for game players, Bartle's model lacks empirical testing, and the player types have not been shown to be independent or validated (Johnson and Gardner 2010, Dixon 2011, Yee 2006).

Yee (2006) offered an empirically derived model for player motivation following Bartle's qualitative discussion, predicated on achievement, social interaction, and immersion. However, Yee developed this model solely through analysis of expert players of massively multiplayer online games, and the model has as yet not been generalised (Bateman et al. 2011).

Drawing on the Myers-Briggs Type Indicator (MBTI), a psychometric measure developed by Cook Briggs and Isabel Briggs Myers (Briggs 1962), Bateman and Boon (2006) developed a more generalisable model of players known as Demographic Game Design Model 1 (DGD1) in which players are categorized into Conqueror, Manager, Wanderer, or Participant types, although these play styles do not appear to relate to Bartle's player types. DGD1 is based on MBTI, and given the inconsistent results from MBTI as a type theory (Boyle 1995) as opposed to the more consistent ones from the five factor model as a trait theory, Bateman et al. (2011) argue that DGD1 suffers from the same limitations as MBTI. The updated version of DGD1, Demographic Game Design Model 2 (DGD2), proposed by Bateman et al. (2011) has yet to be used by other researchers.

The attempt by Ferro et al. (2013) to unify personality and player typologies to create a personalised gamified system, also based on the works of Bartle (2006) and Nacke et al. (2011), resulted in a player type model which included the following types: Dominant, Objectivist, Humanist, Inquisitive, and Creative. However, as with DGD2, the work remains theoretical and has not been validated empirically.

Within the context of this study, one model requirement is generalisability and impact in predicting behaviour, as well as existence of empirical studies based on relationships with models of social interaction and motivation. None of the models within game research satisfied these requirements, and so focus shifted to one of the most widely utilised models of personality (Digman 1990, O'Connor 2002), the five factor model (FFM) proposed as *trait theory* by McCrae and Costa (1987).

The five factor model (FFM) offers a parsimonious taxonomy of personality traits composed of five factors: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (McCrae and Oliver 1992), with empirical research showing consistency in measurements associated with these factors when compared across interviews, self-reports, and observations and encompassing a wide range of individuals of differing ages and cultures (Barrick et al. 1993, Schacter et al. 2009). The personality factors appear to be heritable (Jang et al. 1998) and unaffected by social influence from parents or peers (Riemann et al. 1997). The acronym OCEAN is employed to represent the five domains of personality taxonomies (Costa and McCrae 1992):

Openness (or Openness to Experience) – The individual's tendency toward intellectual curiosity, imagination, creativity, perceptiveness, and joy of variety. A low score tends to describe individuals having a more conventional and conservative outlook.

Conscientiousness – The individual's tendency toward orderliness, decisiveness-consistency, punctuality, and reliability. A low score suggests a subject who is less precise and directed when working towards a goal.

Extraversion – The individual's tendency toward sociability, assertiveness, adventurousness, and active engagement. A low score suggests a more reserved and independent subject.

Agreeableness – The individual's tendency to be sympathetic to others, be cooperative, demonstrate modesty-humility, interpersonal warmth-affection, and generosity. A low score suggests an individual that is competitive, egocentric, and sceptical of others' intentions.

Neuroticism (Emotional Stability) – The individual's levels of emotional reaction to events, emotional security, irritability, and moodiness. High scores tend to describe individuals who experience fear, sadness, embarrassment, disgust, anger, and guilt more than those with relatively lower scores, who are usually even-tempered, relaxed, and calm.

Although FFM is sometimes presented as a universal measure of personality (McCrae and Costa 1997), studies have pointed to such limitations as the following: reliance on self-report instrumentation (McAdams 1992) and issues with representation in some indigenous lexical structures, particularly some Eastern languages, leading to inconsistent outcomes (Saucier and Goldberg 2001).

2.3.1 Five Factor Model of Personality in Games Research

Heinzen et al. (2015) attempted to map the personality types of Bartle (2006) and Yee (2006) to FFM; however, whilst a conceptual overlap has been presented, the study lacked empirical validation.

A relatively small number of studies have applied FFM to game player preferences or satisfaction. Opponents of FFM in game studies, citing lack of relevance or effectiveness, include Orij (2014) and Bateman et al. (2011). However, a re-examination of the original studies used to derive the assertion of "failure of FFM in game studies" by Bateman et al. (2011) discount opposition to exploring FFM with games, as significant results have been observed through the application of FFM in game analysis (Teng 2009). One possible explanation for the dismissal of FFM in game studies by Bateman et al. (2011) and Orij (2014) may, in part, be related to inconsistent instrumentation used within the study's experiment and also to a desire to establish new personality models exclusively for the games domain. New models proposed by Bateman et al. (2011) have yet to be adopted or validated by other researchers.

Analysing the moderating effect of FFM traits with respect to gaming genre preferences (n=450), Zammitto (2010) found significant results between personality traits and game genres. The study showed FFM personality traits as moderating player preferences with respect to the following game genres: Action Shooting, Action No Shooting, Action Fighting, Sports, Simulation Vehicle, Simulation Artificial Intelligence, Adventure, and Puzzle. Teng's (2009) study of online game player (n=114) personality and real-life needs fulfilment indicated a range of FFM traits correlated with fulfilling individual needs for achievement, autonomy, affiliation, and dominance. Seok and DaCosta (2015) investigated FFM traits as moderating mobile game play (N=1997) and found significant results, although with weak overall strength of association.

2.3.2 Five Factor Model of Personality in Gamification Research

Published studies on the correlation between the FFM and game research are extremely limited, and published empirical studies on FFM and gamification research number less than a handful. Researchers in this domain include Codish and Ravid (2014), who studied the impact of various FFM traits on perceived playfulness of gamification in an academic context as expressed through points, leaderboards, and badges, with 102 students (n=102). A significant moderating effect was demonstrated, although effect sizes were small, using partial least square modelling with Q^2= 0.12 for leaderboards and Q^2= 0.14 for rewards (Codish and Ravid 2014).

Karanam et al. (2014) examined a range of game elements (including points, rewards, feedback, challenges, and progress) of a gamification system for daily habits (in categories including mood, health, fitness, diet, and sleep or mood) for graduate students and found correlations between some personality traits and specific reward and challenge preferences. However the small sample size (n=36) with the large variety of potential factors makes generalisation based on these findings difficult (Karanam et al. 2014).

Beyond those of games and gamification, the FFM has been applied in many domains relevant to this study, where moderating impact on outcomes and performance are critical, and the development of studies linking specific personality traits to specific performance criteria, including academic performance (Gray and Watson 2002, Taggar 2002) and commercial (i.e., job) performance (Barrick et al. 2001, George and Zhou 2001), has produced a deeper understanding of personality-performance relationships.

2.3.3 Personality and Performance

Although a wide range of studies have focused on the relationship between personality and performance, for purposes of this study, evidence from the educational and business domains were examined and explored together. Two reasons underlie this decision: First, gamification has the largest applications potential in impacting academic and work performance, and, second, the experimental instrument was developed with both education and business applications in mind.

A large body of literature provides compelling evidence as to the existence of a relationship between personality traits as measured by FFM and job performance (Barrick and Mount 1991, Hurtz and Donovan 2000, Peeters et al. 2006). An equally

significant body of literature evidences the existence of relationships between performance within an academic setting and personality as measured by FFM (Poropat 2009). Table 2.4 below provides an overview of correlations found in the most cited meta-analyses by researchers in the personality-performance domain. Effect sizes reported tend to be small to moderate as defined by Cohen (1998) and Hopkins (1997), in the region of R^2=.02 to R^2=.25.

Table 2.4: Correlations in FFM Traits from Most Cited Meta-Analyses.

	Task Performance	Job Performance	Academic Performance	Professional Performance
Openness	.06**	.07*	.12***	.13**
Conscientiousness	.24*	.24**	.24***	.29**
Extraversion	.09**	.15**	.17***	.11**
Agreeableness	.12**	.11*	.07***	.35*
Neuroticism	-.14*	-.15**	-.02***	-.09**
Meta-Analysis Study Reference	Hurtz, & Donovan (2000)	Barrick et al. (2001)	Poropat (2009)	Peeters et al. (2006)

Reported R^2 values based on true-score correlations.

* p <.05, ** p < .01, *** p <.001

2.3.3.1 Moderating Factors of Personality and Performance

Empirical studies showing the relationship between personality traits and performance are generally moderated by two types of variables, context and task type. *Context* defines the type of situation under which the activity is classified. For example, jobs that require social interaction (Barrick and Mount 1993) provide a context where personalities exhibiting high measures of sociability, i.e., high extraversion, are more likely to perform better.

Other contextual factors include team composition, interpersonal orientation between team members, and relative performance of each individual team member, all of which affect overall team effectiveness and performance (Ross et al. 2003). Task type or nature of the task also moderates effects of personality, and, depending on the complexity and stimulating nature of the tasks to be performed, personality traits may or may not affect performance of the task (Eysenck 1982).

Autonomy as it relates to the individual's volition and control over the context has also been observed as still another contributing, moderating factor in personality-performance (Barrick and Mount 1993, Gellatly and Irving 2001). Individuals with relatively higher scores in Conscientiousness, who are responsible, persevering, dependable, and achievement oriented, and those with high scores in Extraversion, who are assertive, sociable, and outgoing, perform better in work situations marked by relatively high autonomy, i.e., work that offers a high level of discretion in the way it is carried out. Conversely those with lower scores in Agreeableness, who are less soft-hearted and cooperative, performed better in jobs with relatively higher autonomy (Barrick and Mount 1993, Lee et al. 1990). The inverse relationship between Agreeableness and performance as moderated by autonomy may be explained as a conflict between the individual's need to follow rules and lack of strict rules to follow within the job. Individuals rated relatively high on Agreeableness generally prefer to follow rules and be cooperative and therefore seek jobs with a high level of structure, little uncertainty, and hence lower autonomy (Barrick and Mount 1993).

2.4 Hypothesis Formation

This section explores the interaction between interdependence and performance as moderated by personality, leading to the formulation of the main hypotheses for testing in the empirical study.

2.4.1 Openness, Interdependence, Performance

Meta-analyses of studies exploring performance and Openness within a variety of jobs have consistently found this trait to be the FFM factor least correlated with performance, with relatively small positive effects on performance of between R^2=.01 to .09 (Barrick et al. 2001, Hurtz and Donovan 2000). Within academic performance, numerous researchers have demonstrated a moderate positive relationship between Openness and academic achievement (Poropat 2009). However, whilst Openness appears to be positively related to learning motivation (Tempelaar et al. 2007, Vermetten et al. 2001), critical thinking (Bidjerano and Dai 2007), and intelligence (R^2=.15), the overall moderating effect on performance remains relatively small with R^2<.09.

Openness has also been found to negatively moderate the desire for individuals to engage in collective activity, whether cooperatively or competitively (O'Reilly et al. 1991, Koestner and Losier 1996). However, Driskell et al. (2006) found openness to play a positive role within cooperative group activity as a promoter of teamwork, and Musson et al. (2004) found a negative relationship (r=-.21, p<0.05) between competitiveness and openness while examining the desire for individuals to succeed in competitive interpersonal situations.

More recently, Morawetz et al. (2014) utilised functional magnetic resonance imaging (fMRI) to measure differences in brain activity and found that individuals with relatively higher levels of Openness displayed significantly elevated levels of brain activity in 'winning' conditions in contrast to those with relatively lower levels. Whilst it may seem to contradict earlier research by Koestner and Losier (1996), who found individuals with higher degrees of openness prefer individualistic rather than group activity, the neuropsychology findings of Morawetz et al. (2014) provide strong evidence that those with higher Openness levels are affected to a significantly greater degree in interdependent situations. As a result, individuals with higher Openness levels may show aversion toward group activity as a result of being more affected by the

presence of others and hence prefer neutral or individualistic structures. However, if interdependence is not subject to choice, those with higher Openness levels may strive for more positive outcomes for the team within both a cooperative and a competitive setting.

Openness has been found to be the strongest predictor of positive beliefs and optimistic attitudes toward cultural and personal differences within the FFM (Ekehammar and Akrami 2003). Matsumoto and Hwang (2011) found Openness significant in intercultural interaction within teams when examining cooperation and competition within a prisoner's dilemma game. Higher levels of Openness contributed to greater cooperation, as individuals having relatively higher levels more readily engaged with those from dissimilar cultural backgrounds.

Thus, the hypotheses for the relationship between interdependence and performance as moderated by Openness may be stated as follows:

> H_{O1}: Individuals with relatively higher Openness scores exhibit higher performance levels than those with relatively lower Openness scores under cooperative conditions.
>
> H_{O2}: Individuals with relatively lower Openness scores exhibit higher performance levels than those with relatively higher Openness scores under competitive conditions.

Additionally, the seeming aversion to interdependence (Koestner and Losier 1996) of those with high Openness levels leads to the following hypothesis:

> H_{O3}: Individuals with relatively higher Openness scores exhibit higher performance levels in independent conditions than in interdependent conditions.

2.4.2 Conscientiousness, Interdependence, and Performance

The meta-analysis by Barrick and Mount (1991) suggested that Conscientiousness is associated with the strongest overall job performance with $\rho = .2$. Indeed, a further meta-analysis by Barrick et al. (2001) found the same strong relationship and pointed out that imagining an individual marked by carelessness, irresponsibility, impulsiveness, and low scores in achievement striving, all factors contributing to a low Conscientiousness score, doing well in a job was difficult. In addition, a meta-analysis by Poropat (2009)

also showed Conscientiousness to be the strongest moderator for academic performance, even after controlling for intelligence.

Despite being the most highly correlated factor with performance, Tett et al. (1991) claimed the average corrected correlation of Conscientiousness with job performance across various occupations remains relatively modest at $\rho=.22$, possibly attributable to personality-performance correlations having non-linear relationships, a factor that is not typically tested for (Taggar and Parkinson 2007). Additionally contributing to the relatively modest value of ρ could be differences in computation attributable to use of different scales and criterion dimensions across the range of meta-analyses (Hurtz and Donovan 2000).

Highly conscientious individuals tend to perform at a relatively higher level and to experience fewer task conflicts in highly cooperative conditions than when performing tasks that require low levels of cooperation. On the other hand, those with low Conscientiousness levels seem unaffected by level of cooperation associated with an activity (Steinke 2011). Antonioni and Park (2001) found similar evidence for teams, with high levels of Conscientiousness leading to better interpersonal relationships, less conflict, and increased performance amongst team members. This trait is relevant in virtual communication where the organised nature of individuals with high Conscientiousness enables the creation of more comprehensive communication strategies. As these individuals exhibit a higher level of tenacity and persist when faced with difficulties, a number of researchers have noted Conscientiousness as a significant predictor of leaders in virtual teams and, at high levels, a factor in improving virtual team performance (Kozlowski and Bell 2003, Cogliser et al. 2012).

The moderating effect of Conscientiousness appears to be significantly moderated by such contextual factors as social loafing as well as task type (Schippers 2014, Steinke 2011, Hough 1992). High Conscientiousness appears to lead to higher performance levels where opportunities for social loafing are limited or nonexistent. In contexts with the opportunity for social loafing, individuals with high Conscientiousness appear to perform worse.

Task type also affects the moderating effect of Conscientiousness, with evidence showing that tasks requiring high levels of decision making or creativity are negatively related to this trait whereas tasks that require more planning and coordination appear to be positively related (Neuman and Wright 1999, Barry and Stewart 1997). In contrast,

Judge et al. (2014) found that Conscientiousness positively moderates creativity. These contrasting findings may be due to the nature of creativity as interpreted in different situations; for instance, creativity within an artistic setting may differ from creativity within a scientific setting.

Exploring for moderating factors for Conscientiousness among hypercompetition, personal development, and cooperation, Ross et al. (2013) found a modest positive effect ($\rho = .13$, $p < .05$) for personal development competition. However, the study did not report any significant differences between cooperation and hypercompetition. The positive correlation between Conscientiousness and performance under personal development competition may be explained as the trait's making an individual more careful, thorough, and responsible, thus causing personal responsibility and personal development to appeal more to those scoring higher in this trait.

These findings and incorporation of social loafing yields the following hypotheses:

> H_{C1}: Within contexts providing greater opportunities for social loafing, individuals with relatively lower Conscientiousness scores exhibit higher performance levels than those with relatively higher Conscientiousness scores under cooperative conditions.
>
> H_{C2}: Within contexts providing greater opportunities for social loafing, individuals with relatively higher Conscientiousness scores exhibit higher performance levels than those with relatively lower Conscientiousness scores under competitive conditions.

2.4.3 Extraversion, Interdependence, and Performance

Several researchers have found Extraversion to be positively correlated with both enjoyment (Graziano et al. 1997, Kirkcaldy and Furnham 1991) and performance (Bentea and Anghelache 2012, Chen et al. 2011) under competitive structures (Judge and Zapata 2014). However, the positive relationship between Extraversion and job performance appears more prominently in job performances where interpersonal interaction is a key aspect of the contextual task (Costa et al. 1984, Mount et al. 1998, Van Scotter and Motowidlo 1996). Thus, higher levels of Extraversion have been demonstrated to lead to higher levels of group participation (Littlepage et al. 1995).

Graziano et al. (1997) observed that the propensity for an individual to find competitive games more likeable and interesting than cooperative games increased as Extraversion

increased. Extraverts are more likely to view social encounters as positive and so tend to gravitate toward situations where strong social skills are needed, for instance, in competitive social occupations (Gallagher 1990). Since extraverted individuals seek sensation and interpersonal contact (Costa and McCrae 1985), it would follow that such individuals would initiate actions or prefer activities that offer excitement. Competition increases the potential for stimulating activities based on interpersonal relationships that involve risk taking for gain and hence appeals to a greater degree to those higher in Extraversion (Bentea and Anghelache 2012).

Stewart (1996) also found Extraversion-performance to be moderated by interdependence, specifically observing that high-Extraversion individuals tend to excel where performance comparisons are made explicit. This finding suggests that leaderboards or other appraisal methods that contrast one's performance with another's may be effective for highly extraverted individuals within a team.

Stewart's (1996) findings were further supported by those of Bentea and Anghelache (2012), who demonstrated significant effect sizes with relation to Extraversion between competitive and cooperative cognitive tasks. Extraverted individuals tended to perform better under competitive and individualistic conditions compared to those with low Extraversion, who performed less well under competitive or individualistic conditions. Low levels of Extraversion in individuals seem to contribute to greater performance under cooperative conditions. Individuals falling in the approximate middle of the Extraversion scale tended to perform as badly as those with low levels of Extraversion under individualistic conditions. Moreover, the performance of these individuals fell almost half way between that of high extroverts and low extroverts under cooperation (Brentea and Anghelache 2012).

Extraversion has been found to be one of the strongest predictors of achievement motivation under both personal development competition ($r=.356$, $p<.001$) and hypercompetition ($r=.148$, $p<.01$) amongst all the FFM traits (Ross et al. 2003). The subtle distinguishing elements of Extraversion for those who perform higher in hypercompetition were found to relate to lack of interpersonal warmth and positive affect toward other participants. Individuals who either avoid or ignore positive interpersonal connections perform better in hypercompetition by focusing on 'winning for the sake of winning' and avoid being negatively affected by concern for others. Conversely, those who preferred personal development competition tended to emphasise the importance of personal achievement and positive interpersonal

connections. These findings support earlier findings of Ryckman et al. (1997) that distinguished the correlation of Extraversion with hypercompetition and personal development competition.

The findings reported above yielded the following hypotheses:

> H_{E1}: Individuals with relatively lower Extraversion scores exhibit higher performance levels than those with relatively higher Extraversion scores under cooperative conditions.
>
> H_{E2}: Individuals with relatively higher Extraversion scores exhibit higher performance levels than those with relatively lower Extraversion scores under competitive conditions.

2.4.4 Agreeableness, Interdependence, and Performance

In response to the Barrick et al. (2001) meta-analysis indicating an overall limited relationship of Agreeableness and performance, Johnson (2003) noted that Agreeableness may aid performance in some jobs but be a limiting factor in others. To be sure, Mount et al. (1988) also found that Agreeableness was a valid positive predictor of performance (r=.35) only in jobs involving cooperation or dyadic interaction. Van Scotter and Motowidlo (1996) provided further supporting evidence in a study demonstrating a greater positive relationship between Agreeableness and performance in jobs involving interpersonal interaction but little effect on performance in jobs with limited to no social dimension. Chiaburu et al. (2011) provided still more evidence in support of Agreeableness and prosocial work behaviour.

Individuals with high scores in Agreeableness tend to maintain positive interpersonal relationships with others (Barrick et al. 2002). A meta-analysis by Peeters et al. (2006) provided further evidence that teams with higher overall Agreeableness performed significantly better in cooperative situations.

Agreeableness is characterised by the individual's preference for cooperation over competition (McCrae and Costa 1987, Liao and Chuang 2004). The level of an individual's Agreeableness within a team has shown to be one of the strongest predictors of group performance, as teams composed of individuals exhibiting higher levels of Agreeableness consistently perform better under cooperative structures than under competitive structures (Bell 2007, Beersma et al. 2002, Schippers 2014) as more agreeable individuals not only tend to be better at interpersonal interactions (Hurtz and

Donovan 2000) but are also more motivated toward positive social interaction, aimed at reducing within-group competition, reducing conflict within the team, and promoting greater coordination.

Ross et al. (2003) study (n=249) demonstrated Agreeableness to be the most significant of the FFM traits as a moderator for an achievement orientation toward competition and cooperation. Hypercompetitive structures were shown to correlate negatively to a significant degree with Agreeableness (r=-.584, p<.001) whereas cooperative structures exhibited a positive correlation (r=.346, p<.001). Personal development competition, however, was not found to be significant. Schipper (2014) confirmed the observation that Agreeableness is a significant predictor of performance in cooperative activities, with teams composed of individuals with higher levels of Agreeableness outperforming those composed of individuals exhibiting lower Agreeableness levels.

In a longitudinal study on the impact of Agreeableness on team performance, Bradley (2013) found that affect-based states such as team cohesion tend to need longer to develop than behavioural interactions such as communication. Furthermore, the virtual or physical nature of interpersonal interaction was shown to impact the effect of Agreeableness, which positively improved physical face-to-face interaction between team members, incorporating tone of voice, inflection, and body language. Virtual interaction through texts and email, on the other hand, drastically reduced the richness of interpersonal communication and so was shown to diminish any impact Agreeableness could exert on communication. The finding that virtual communication in teams reduces the perceived levels of Extraversion and Agreeableness in their team mates is also consistent with earlier findings of Vignovic and Thompson (2010) and general findings on impact of virtual team dynamics (Becker-Beck et al. 2005).

Bradley's (2013) findings that Agreeableness leads to improved performance only where team members interact face-to-face as opposed to virtually and that low-agreeable teams performed better than teams with high Agreeableness when communication is restricted to virtual means are significant in exploring this trait as a moderator of team performance, particularly in theoretical models incorporating virtual teams such as social presence theory, and has managerial implications in the real-world. It underlines the need for face-to-face interaction, particularly initially to improve team cohesion, increase trust, and enable the Agreeableness trait to promote harmony and social integration within the team (Bradley et al. 2013).

Graziano et al. (1997) explored competitiveness as a mediator of the link between Agreeableness and group performance, noting two possible ways in which the Agreeableness trait could impact performance under cooperation or competition. Firstly the expected social relation may be perceived as competitive prior to any interaction taking place, and secondly dispositionally competitive people are more prone to see others as competitive and so evoke competitive behaviour from others (Kelley and Stahelski 1970, Graziano et al. 1997).

The findings from the existing literature reported above yield the following hypotheses:

> H_{A1}: Individuals with relatively higher Agreeableness scores exhibit higher performance levels that those with relatively lower Agreeableness scores under cooperative conditions.
>
> H_{A2}: Individuals with relatively lower Agreeableness scores exhibit higher performance levels than those with relatively higher Agreeableness scores under competitive conditions.

2.4.5 Neuroticism, Interdependence, and Performance

Neuroticism reflects an individual's emotional instability. Individuals scoring high on Neuroticism tend to avoid setting goals and may display greater insecurity, hostility toward others, moodiness, and anxiety. These characteristics negatively impact individual motivation and can also create a negative atmosphere within a team, negatively affecting group performance (Judge and Ilies 2002, Mount et al. 1998). But whilst elements of Neuroticism such as guilt and dissatisfaction may help promote self-improvement, other elements (e.g., impulsiveness and anger) may result in self-destructive behaviour and negatively impact motivation (Shebaya 2011).

The link between performance and Neuroticism has consistently been shown to be negative. A meta-analysis by Barrick et al. (2001) found a correlation of ρ =-.15 between job performance and Neuroticism. Similar effect sizes were found by others, including Hurz and Donovan (2000) (ρ=-0.14) and Mount et al. (1998) (ρ=-.15 to ρ=-.25) with respect to job performance. The level of interaction between Neuroticism and performance is similar to the effects of Agreeableness. However, the moderating impact of Neuroticism on academic performance appears to be minor, with effect sizes of ρ=-.02 reported by Poropat (2009).

Neuroticism has been found to negatively correlate with personal development competition (Ryckman et al. 1990) and positively with hypercompetition (r=.23, p<.001) (Ross et al. 2003) and (r=.54, p<.01) (Chen et al. 2011, Ryckman et al. 1996), respectively. Although the Ross et al. (2003) study did not find a correlation between a global measure of Neuroticism and personal development competition, the study did find a moderate negative correlation between specific facets of Neuroticism, specifically neuroticism/depression (r=-.31, p<.001) and neuroticism/self-consciousness (r=-.24, p<.001). Indeed it seems that individuals with high levels of Neuroticism enjoy the sense of superiority that hypercompetition offers (Ross et al. 2003).

Moberg (2001) and Antonioni (1998) both found Neuroticism to negatively correlate with confrontation, suggesting that individuals with high Neuroticism may seek to avoid situations where confrontation could occur such as in hypercompetition. However hypercompetition does not necessarily imply conflict but rather a zero-sum interaction which incorporates little to no concern for possible harm to others (Ryckman et al. 1997). Thus, whilst Neuroticism contributes positively to the individual avoiding conflict, since it increases sensitivity to perceived threats (Derryberry and Reed 1998), it appears to fuel the desire to achieve dominance and hence positively moderates hypercompetition (Fletcher and Nusbaum 2008).

Fang (2002) found Neuroticism to negatively affect knowledge sharing with others in a team, thereby negatively impacting team cooperation. Although some studies have not found a significant correlation between Neuroticism and cooperation (Ross et al. 2003), others, including Chen et al. (2011) (r=-.40, p<.01), have found a strong negative correlation. Lonnqvist et al. (2011) found individuals with low Neuroticism more cooperative in a prisoner's dilemma game, suggesting a negative correlation between cooperation and Neuroticism. Given the similarities between personal development competition and cooperation, the findings of Lonnqvit et al. (2011) appears in line with the earlier work of Ryckman et al. (1990) with respect to personal development competition.

All forms of social interaction may carry an aspect of social comparison, whether competition, cooperation, or even a simple discussion (Brickman and Bulman 1977 as quoted in Buunk et al. 2005). Buunk et al. (2005) found a significant correlation between Neuroticism and a social comparison orientation (r=.3, p<.001), suggesting that the trait is a positive predictor of social comparison.

The findings above yielded the following hypotheses:

> H_{N1}: Individuals with relatively lower Neuroticism scores exhibit higher performance levels than those with relatively higher Neuroticism scores under cooperative settings.
>
> H_{N2}: Individuals with relatively higher Neuroticism scores exhibit higher performance levels than those with relatively lower Neuroticism scores under competitive settings.

2.5 Summary

This chapter has explored the multiple effects of personality and interdependence on individual performance. The evidence that both personality and interdependence moderate performance is itself moderated by context.

The literature review chapter identified and presented a number of hypotheses to be tested empirically. By way of summary, the hypotheses can be restated as follows:

H_{O1}: Individuals with relatively higher Openness scores exhibit higher performance levels than those with relatively lower Openness scores under cooperative conditions.

H_{O2}: Individuals with relatively lower Openness scores exhibit higher performance levels than those with relatively higher Openness scores under competitive conditions.

H_{O3}: Individuals with relatively higher Openness scores exhibit higher performance levels in independent conditions than in interdependent conditions.

H_{C1}: Within contexts providing greater opportunities for social loafing, individuals with relatively lower Conscientiousness scores exhibit higher performance levels than those with relatively higher Conscientiousness scores under cooperative conditions.

H_{C2}: Within contexts providing greater opportunities for social loafing, individuals with relatively higher Conscientiousness scores exhibit higher performance levels than those with relatively lower Conscientiousness scores under competitive conditions.

H_{E1}: Individuals with relatively lower Extraversion scores exhibit higher performance levels than those with relatively higher Extraversion scores under cooperative conditions.

H_{E2}: Individuals with relatively higher Extraversion scores exhibit higher performance levels than those with relatively lower Extraversion scores under competitive conditions.

H_{A1}: Individuals with relatively higher Agreeableness scores exhibit higher performance levels that those with relatively lower Agreeableness scores under cooperative conditions.

H_{A2}: Individuals with relatively lower Agreeableness scores exhibit higher performance levels than those with relatively higher Agreeableness scores under competitive conditions.

H_{N1}: Individuals with relatively lower Neuroticism scores exhibit higher performance levels than those with relatively higher Neuroticism scores under cooperative settings.

H_{N2}: Individuals with relatively higher Neuroticism scores exhibit higher performance levels than those with relatively lower Neuroticism scores under competitive settings.

In the next chapter, these hypotheses are scoped into the context of this thesis's experimental instrument.

3 RESEARCH METHODOLOGY

This chapter focuses on discussing the study's research methodology and experimental design. It describes the process whereby the experimental instrument, *StarQuest,* was developed, the results of the formative evaluation carried out during the pilot study, and the methodologies used in the quantitative evaluation of the experimental results.

Thus, in greater detail, the objectives of this chapter are as follows:

- Justify the study's research philosophy, paradigm, and strategy
- Describe the method used for designing the experimental instrument
- Discuss the operationalisation of the dependent and independent variables and their validity through the experimental instrument
- Describe the data collection, the study's participants, and related issues concerning ethics
- Highlight the analytical methods utilised
- Discuss the main sources of error and limitations of the study

3.1 Research Philosophy and Strategy

The focus of this research is measuring the impact of game dynamics on outcomes within a social setting by collecting data regarding the behaviour and actions of individual participants and groups in an online experiment. Thus, the theoretical underpinning of this research is deterministic philosophy in which cause determines outcome. This assumption of cause and effect constitutes the scientific method as

expressed in quantitative research, positivist/postpositivist research, empirical science, and postpositivism (Cresswell 2003).

In this study, quantitative research methods are used to analyse measurements related to the characteristics, traits, and attributes of social entities via statistical and computational techniques (Neuman 2006). Positivism is a philosophy of science where information is derived from mathematical and quantitative reports of sensory experience (Macionis and Gerber 2011).

The postpositivist tradition draws on the work of 19th century writers who rejected the positivist view that the researcher and a subject of the research are independent of one another (Smith, 1983). More recently, Phillips and Burbules (2000) describe postpositivism as not simply challenging the traditional notion of absolute truth of knowledge but as necessitating experiments to examine the causes that influence outcomes within a reductionist framework. Whilst postpositivism generally retains the idea of objective truth, it also acknowledges that theories, background, knowledge, and even values of the researcher may influence what is observed (Robson 1993).

The reductionist position aims to explain phenomena in terms of relationships between more fundamental elements and can be applied to the behavioural sciences, where a complex system may be described through a hierarchy of constructs, each of which is described with reference to simpler objects (Dawkins 1996). Thus by reducing the overarching ideas into a small, discrete set of ideas and measurable variables, hypotheses can be constructed to address research questions in order to examine the objective reality present “in the world” (Cresswell 2003).

By drawing on the works of Phillips and Burbules (2000), Cresswell (2003) summarises the key positions of postpositivism as follows:

- Absolute truth can never be found, as knowledge is formed as an opinion based on incomplete information. Hence, evidence merely supports hypotheses rather than proving them conclusively.
- The research process is redefining or refining claims based on receipt of information.
- Knowledge is formed through rational consideration of data and evidence.

- Research seeks to describe the causal relationship among variables posed in terms of questions or hypotheses in order to arrive at statements perceived to be true.
- It is essential that the inquiry be carried out objectively, examining methods and conclusions for bias and adhering to high levels of validity and reliability.

Within the context of this study, the researcher has adopted a postpositivist stance with respect to analysing its objectives. Thus, study measures, which are quantitative, are made objectively, utilizing methods that are as valid and reliable as possible, and study results are observed with utmost rationality, whilst also acknowledging that the researcher's considerable background knowledge on the study's topic and values may influence the inferences derived from the observations and claims made.

3.1.1 Design of Experimental Instrument

This section describes the method used to design and develop the experimental instrument, an online platform called *StarQuest*. Although *StarQuest's* design and development occurred over the course of the researcher's Ph.D. programme, its development is presented as part of the study's contribution, as modifications to the platform were required to create the experimental conditions needed to conduct the empirical study described herein.

3.1.1.1 Embedding the Experimental Instrument in Participants' Lives

This study relies on collecting data from participants through their interaction with modified versions of the online platform *StarQuest*, which facilitates collaboration between small groups of participants in the collection, creation, and sharing of digital artefacts. Functionality was added to the platform to enable participants to be introduced to the study, complete and submit informed consents, and access the experiment's different conditions through manipulations of the platform's user interface. Other functionality added was the ability to measure, collect, and store data on participant responses for later analysis. Furthermore, the experimental instrument (i.e., *StarQuest*) provides end-user utility such that participant activities undertaken within the platform form a part of their existing daily activities within the context of their studies, rather than being carried out in an isolated setting such as a lab.

Collecting information through online platforms, whether through surveys, remote user testing (Andreasen 2007), or online experiments (Fogg et al. 2001), reduces costs of collecting user input. When the topic being studied concerns the use of online platforms, as was the case in this study, data collection in the field, accomplished by observing participant behaviour as part of their natural activities as opposed to in laboratory conditions, is possible. Horton et al. (2011), examining the validity of online experiments in contrast to that of laboratory and field experiments, conducted experiments which confirmed that the external validity of participants' responses to priming, pro-social preferences, and decisions in response to how a decision-problem was framed equalled or exceeded that of traditional laboratory methods.

The literature review reported a considerable disparity between the motivational increases reported by the users of gamified systems and actual increases in performance. Since users of gamified systems are situated in a real world context, in this instance within a learning environment, an experimental approach focusing on the users' actual requirements and carried out in-situ may provide a higher level of validity for performance impacts (Horton et al. 2011) than experiments carried out in artificial, laboratory environments. In a meta-analysis of experimental settings, Levitt and List (2006) conclude that the dichotomy sometimes drawn between data collected in lab experiments and generated in natural settings is not always warranted. However, obtaining deep structural parameters such as those sought by this research study is more likely by embedding the study instrument in a natural environment than in isolated laboratory conditions.

3.1.1.2 Design via a User-Centric Design Methodology

As the experimental instrument is embedded in the participants' natural use environment, it must provide an acceptable level of utility compatible with end-user requirements and activities. Thus, understanding user requirements formed a critical part of the process of developing the instrument, and therefore a user-centred design process focused on the end users' needs, wants, and limitations was adopted (Kramer et al. 2000). Specifically, the methodologies formalised in the ISO standard on 'Human-centred design for interactive systems' (ISO 9241-210, 2010) were followed.

Focusing on user needs and requirements to accomplish human-centred design was accomplished by way of the following: basing design on an explicit understanding of users' tasks and environments, involving users throughout the design and development

processes, and employing user-centred evaluation in the design and refinement processes in an iterative process.

A wide range of techniques exist to elicit understanding of users' tasks and environments in order to derive requirements, including such traditional techniques as surveys, interviews, and analyses of existing processes and systems; group elicitation via brainstorming and focus groups; prototyping; model-driven techniques utilizing goal- and scenario-based methods; cognitive techniques such as protocol analysis, laddering, card sorting, and repertory grids; and contextual techniques using ethnographical, ethnomethodological, and conversation analysis (Nuseibeh and Easterbrook 2000).

3.1.2 Requirements, Design, Formative Evaluation, and Pilot Trials

As this study utilised a human-centric design process, group elicitation together with prototyping and scenario-based methods were selected as the principal methods of gathering data from end-users. Initial exploratory design workshops involving small groups of end-users were organised in the first year of the study. During the workshop, brainstorming (Bouchard 1971) and creativity techniques including constraint removal (Nguyen et al. 2000, Maiden et al. 2004) and creativity triggers (Goldratt and Whitford 1992, Jones et al. 2008) were utilised in exploring and discussing the range of tasks and activities undertaken by the users. These were carried out with the aim of uncovering gaps within the participants' working or learning environments which an online platform could potentially improve.

Storyboards were used as sequential graphical representations of the systems to depict interactions between the user and the system, highlighting on-screen components for exploring and refining requirements (Andriole 1989). Insights and ideas could thereby be represented through user scenarios, a description of particular tasks performed by a persona in a story form as though it were being actually observed (Holtzblatt and Wood 2004).

Ideas generated within the workshop were then explored through user journey scenarios and represented as storyboards that explored how potential functionalities would meet end-user requirements. These requirements were aggregated and further analysed through a structured iterative design process. The range of requirements were refined to key elements and components which could: (1) satisfy end-users requirements, (2) be

generic enough to have applicability beyond specific users, (3) be within the scope of development and testing within the study timeline, and (4) critically ensure scientific advancement in the domain of the study. Appendix 2 provides extracts of the user scenarios identified, focusing on specific user tasks and interactions which could be supported from which requirements could be developed.

Based on the user requirements, a set of functionalities and user interface designs were sketched out to further solicit user feedback and refinement. The purpose of the platform is to enhance motivational affordance (Jung et al 2010) of small groups to collaborate online. The core functionality of the platform is to provide support for creating and commenting on digital artefacts as manifested through the ability to 'post' new content and 'comment' on previously posted content. This functionality is present in all conditions of the platform and is based on the core requirement for collective gathering and creating of digital artefacts, defined as social creativity, where groups work together to solve problems with the help of digital technologies (Fischer and Wulf 2005). Shneiderman (2000) defines a model for developing user interfaces that support social creativity called GENEX, which offers a four-phase framework for social creativity consisting of (1) collecting, where information is gathered from sources such as the Internet; (2) relating, where peers and mentors are consulted; (3) creating, where composition and evaluations of artefacts occur; and finally (4) donating, where the results are disseminated to a wider audience.

The gamification of the platform was achieved by incorporating a range of game elements including goals, points, status, and leaderboards, developed in such a way as to enable group level manipulation of social interdependence settings for participants. Framing of the interface elements can be manipulated between cooperative and competitive conditions to change the activity's social interdependence in the following mechanisms:

- A points system that provides rewards and penalties for activity or lack of.
- A status system that provides feedback through human-face embodied agents.
- A system to suggest goals to participants together with turn-taking.
- A leaderboard system that can display an individual's effort and position, either in comparison to others (competitive condition) or as a contribution to group effort (cooperative condition).

Formative evaluation with a range of formal and informal assessment techniques was carried out qualitatively and used during the development process to improve students' acceptance of the platform (Scriven 1967, Shepard 2005). The Technology Acceptance Model (TAM) predicts user acceptance based on perceived usefulness and perceived ease of use by users (Davis 1985, Davis 1993). TAM has been validated in a number of studies that have proved its validity in predicted intention to use (Mathieson 1991, Taylor and Todd 1995, Vankatesh and Davis 2000).

Appendix 2 provides a short summary of the range of questions and factors used in informal discussion with end users through the initial feedback and refinement period based on the TAM model (Vankatesh and Davis 2000). The feedback was recorded as notes on the existing designs in order to develop a prototype of the platform for further formative evaluation. The prototype of the platform was then tested by a small group of end-users, leading to further refinement and second rounds of prototyping, incorporating the measurement of the dependent variables and manipulation of the independent variables as described later in this chapter. The outcome of the piloting trials, where a significant number of improvements and changes were discussed and explored were incorporated into the final version of the platform for the full-scale running of the experiment.

3.2 Operationalisation of the Experiment's Constructs

Operationalisation of variables is a process through which the experiment's constructs become quantifiable and measurable, thus enabling the underlying, fuzzy theoretical concepts, collaboration and competition, to be explored in the research and be distinguishable in terms of empirical observation. The concept of operationalisation also relates to measurements of concepts, and, whilst scholars have worked on a wide range of scale and index constructions in areas such as performance, personality, and learning, there is no single perfect method through which to operationalise (Shields and Rangarajan 2013).

The empirical research questions in this study, presented in the introductory chapter as RQ4, RQ5 and RQ6, require measurement of participant personality traits and measurement of the relative difference of participants' performance in *StarQuest* under cooperative, competitive, and control conditions. However, the study was running within existing courses, and the manipulations necessary to create these three conditions and the degrees to which the environment could be controlled required sensitivity to the context.

The ethical issues regarding potential negative consequences in the context of the experiment setting, i.e., undergraduate students in existing courses using *StarQuest,* interdicts the use of the participants' performance in the experimental platform as summative evaluations, i.e., to set grades for a student based on how well he or she performed in *StarQuest*. Additionally, since the *StarQuest* platform does not place limitations on the resources or means available to individual participants, i.e., anyone can find or share any resource or artefact found on the Internet, the experiment incorporated no provision for interdependencies in means. Thus the experiment's manipulations were limited as to which activities or goals and their subsequent outcomes were communicated to the participants within the *StarQuest* platform, used voluntarily by participants and without more far-reaching impact on their grades or outcomes external to the platform.

3.2.1 Independent, Dependent, and Moderator Variables

In summary, the manipulations of conditions and subsequent measurements of performance within *StarQuest* provided a realistic, if imperfect, scenario for cooperation and competition within an educational setting. The actual conditions were aligned with outcome cooperation and outcome competition as described in the literature review, with the caveat that outcomes presented to participants were within *StarQuest* and did not impact the individuals' institutional outcomes.

An empirical experiment can be considered as 'ideal' where manipulation of independent variables can help to establish a causal relationship of these with dependent variables (Grinnell et al. 2008). Grinnell et al. (2008) describe group research designs as those most likely to have external validity, where the findings of the study can be generalised beyond the specific conditions forming the research context (Bortz 2005).

Within this study, the gamification dynamic condition, with levels *competition*, *cooperation*, and *control (neutral)*, acts as the independent variable and manifests through manipulation of *StarQuest's* features that present the required actions, i.e., posting and commenting, and their outcomes in the user interface such as points earned or lost, status gained or lost, and leaderboard position.

Participants' performance generates the dependent variables, measured quantitatively in terms of the number of posts and comments made by participants, the platform's principal user activities, and the length of time spent on the platform.

The moderating variable, which may affect the strength and relation between the dependent and independent variables, is participant personality, as measured by level of FFM personality traits.

All core functionalities of the platform are present and perform in the same way under the three conditions. The difference in conditions are the way in which system functions are presented and activity impacts either the individual or the group.

The diagram in Figure 3.1 below shows the overview of the variables within the study's experiment and the relationship between them. The three independent variables, i.e., the conditions *cooperation*, *competition*, and *control*, constitute the factors in the study's experimental design.

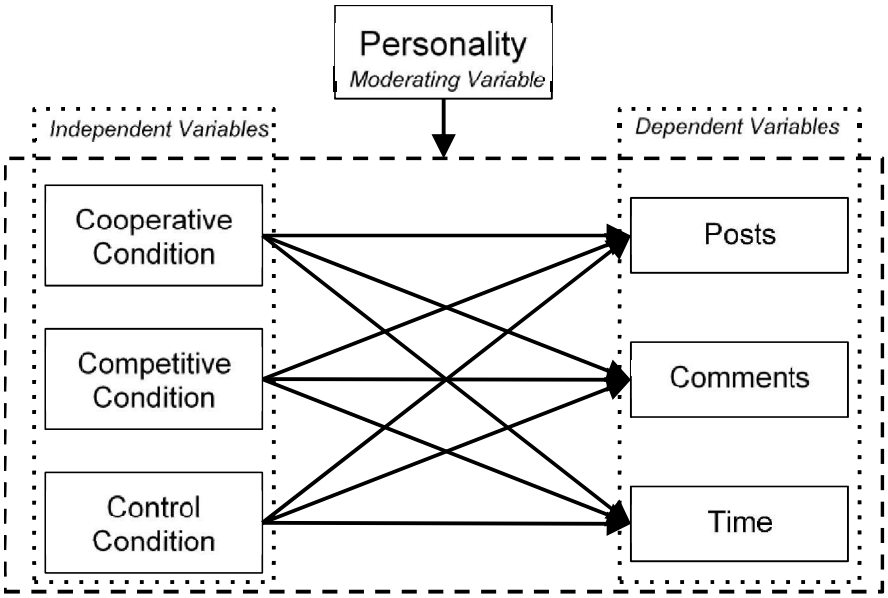


Figure 3.1: Relationship of Study Variables.

3.2.1.1 Independent Variable Manipulation

The independent variable, the interdependent condition, is manifest through gamification of *StarQuest* using points, status, leaderboard, and the framing of actions and outcomes. The functionalities of these are described below. End-user input during the development and piloting period was used to ensure that framing of functions communicated the desired interdependent condition.

Points Mechanism, Reward, and Penalties for Activity or Lack Thereof

StarQuest has been modified to incorporate a point system to provide positive reward points for such activities as contributing posts or commenting and punishment for lack of activity. The point system uses the moniker 'health points.' The use of 'health points' to denote vitality or impact of actions is an established game element and part of affective interfaces for online interaction (Graepel et al. 2004, Nacke and Mandryk 2010, Siang and Rao 2003).

Depending on the experiment's condition, health points are awarded either collectively to the whole group under the cooperative condition or individually under the competitive condition, with the control condition associated with no health points. Health points are awarded based on any activity which impacts the content on the platform, including completing or not completing a goal, posting and commenting, and each type of activity has a certain number of health points associated with it. Table 3.1

summarises the mechanics of health points. Health has a maximum, denoted in a health bar with a maximum of five full hearts possible (see Figure 3.2). Health decays over time; the longer the time interval over which a player is inactive on the platform, the greater the extent to which health decays.

Table 3.1: Manipulation of *StarQuest's* Parameters to Frame Condition.

	Cooperative Condition	**Competitive Condition**	**Control Condition**
Points gained	Posting and commenting by any member of a group earns health points for the group.	Posting and commenting earns health points for the individual and contributes to that individual's leaderboard rank.	There is no point scoring mechanism.
Points lost	Lack of activity by any individual results in loss of health points for the whole group.	Lack of activity results in loss of health points for the individual and loss of position on the leaderboard.	There is no point scoring mechanism.

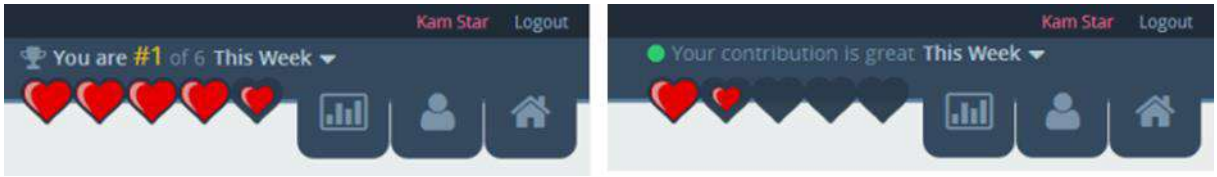


Figure 3.2: *StarQuest's* Health Bar.

Clear goals, control over one's activities, and immediate feedback act as antecedents to achieving flow in online platforms (Chen 2006). The key components of most games are goals, rules, interaction, and rewards (Salen and Zimmerman 2004). Gamification uses similar components through PBL as the principal methods for communicating that the participant has completed the requisite activities (Deterding 2011).

StarQuest provides goals associated with the core activities of finding and creating digital artefacts around specific topics as shown in Figure 3.3. According to SDT (Ryan and Deci 2000), users' autonomy, their ability to determine their own actions, is a critical factor in motivation and well-being; therefore a system in which goals are authoritatively given to the participant by an algorithmic agent may appear counter to autonomy. For this reason, goals are presented through a set of choices to the participants, and the participant has autonomy with respect to choice or may suggest new topics to be explored. This soft approach to goal-setting resulted from the

formative evaluation; initially goals were strict, but in further refinements goals were reframed as suggestions rather than requirements. Goals follow the SMART (Meyer 2003) methodology, which dictates that they be highly specific (e.g., write a paragraph on user interface), measurable (in that their completion is recorded in the system), relevant to the group's activity, attainable, and time bound (goals expire after a certain time if left unchecked).

StarQuest has been modified to provide a turn-taking method for goal completion, where users are asked in turn to complete a task, i.e., posting. User studies have identified turn taking as improving social collaboration in both physical and computer-mediated situations (Shaer et al. 2010, Inkpen et al. 1997). Turn taking is a prevalent game mechanism used in physical games such as chess, digital games such as turn-based strategy, or role-playing games (Bjork and Holopainen 2005, Pinelle et al. 2009). Turn taking evokes social pressure, where expectation of other participants acts as a motivator for action on the part of the individual whose turn it is (Fogg 2002). Turn taking in games often follows a rigid structure involving a strict rotation of turns, such as found in Monopoly. Turns must be taken within a predetermined length of time or are otherwise considered missed. Under the cooperative condition, missing a turn impacts the entire group's health score and happiness status, whereas under the competitive condition, missing a turn only impacts that individual's health score and happiness status.

The manipulations to create the experimental conditions in which the goals and subsequent rewards and punishments are presented to the participant are described below in Table 3.2:

Table 3.2: Manipulation of *StarQuest's* Goal Presentation, Awards, and Punishments to Frame Condition.

	Cooperation Condition	**Competition Condition**	**Control Condition**
Homepage Leaderboard	Collective effort on each subject area.	Displays the name of the individual ranked at first place for each subject area together with the ranking position for the participants.	A total count of posts in each subject area without reference to individual or collective.
Performance Page Leaderboard	Alphabetical list of participants.	Ranked list of participants with individual with highest points at the top.	No performance page is displayed.

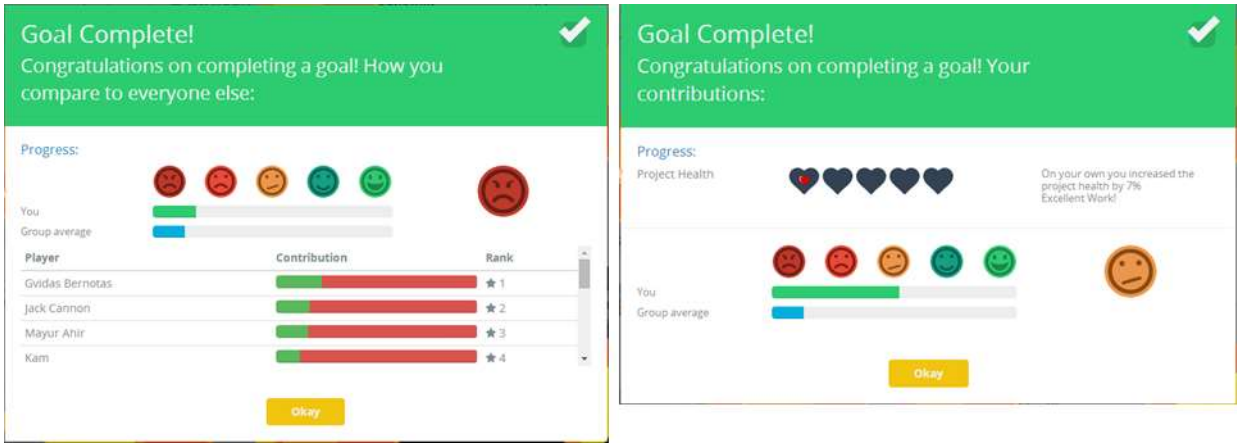


Figure 3.3: *StarQuest's* Goal Framing.

Prompts for Activities

The two key activities in *StarQuest*, posting and commenting, are framed differently depending on condition (i.e., cooperation, competition, control). Table 3.3 below provides the type of prompting together with an example for each condition.

Table 3.3: Manipulation of *StarQuest's* Prompts to Frame Condition.

	Cooperation Condition	**Competition Condition**	**Control Condition**
Prompts for activities in *StarQuest*.	Prompts for activities are framed cooperatively as contributing to the overall group outcomes. The individual is encouraged to complete the action to improve the group's knowledge of a subject area. Completing the activity earns health points, increases the group rank for the subject, and increases group 'happiness' status.	Prompts for activities are framed competitively as contributing to the individual's outcome in comparison with other group members. The individual is encouraged to complete the action to improve his/her own position and ranking. Completing the activity earns health points for the individual, increases his/her leaderboard rank for the subject, and increases his/her individual 'happiness' status.	The activity is not framed. Contributions do not earn health points or change happiness status, as these are not displayed.
Example Prompt for Posting	A list of individuals who have contributed to a particular subject	"You are 1 place behind John on #Dataplay. **Compete with John** by	N/A

(#Dataplay is an example subject area)	area are displayed followed by "… have contributed on #Dataplay. **Help the group with your contribution** by Posting on #Dataplay"	Posting on #Dataplay"	
Display of user effort	"Your contribution to the group is *great* this week"	"You are #2 of 8 this week"	N/A

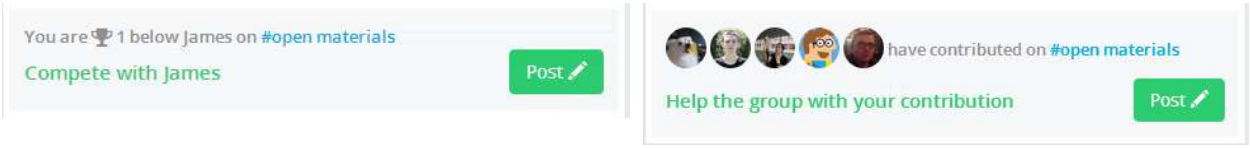


Figure 3.4: Example Manipulation of *StarQuest's* Prompt for Action to Frame Condition.

Leaderboard Mechanisms

StarQuest has been modified to provide two types of leaderboards, a competitive one and a cooperative one. In the cooperative condition, the home page leaderboard displays a list of subject areas for contribution together with the collective effort toward each, whilst the performance page displays a list of all the participants in alphabetical order. In the competitive condition, the home page leaderboards display the name of the individual currently at first position for a given subject area, along with the participant's current leaderboard rank for the specific subject area (see Figure 3.5), whilst the performance page displays a list of all participants in order of their ranks, with individuals with the highest score at the top (see Figure 3.6). There are no leaderboards in the control condition. Table 3.4 summarises the leaderboard element under each condition.

Table 3.4: Manipulation of *StarQuest's* Parameters to Frame Condition.

	Cooperation Condition	**Competition Condition**	**Control Condition**
Homepage Leaderboard	Collective effort on each subject area.	Displays the name of the individual ranked at first place for each subject area together with the ranking position for the participants.	A total count of posts in each subject area without reference to individual or collective.
Performance	Alphabetical	Ranked list of	No performance

Page Leaderboard	list of participants.	participants with individual with highest points at the top.	page is displayed.

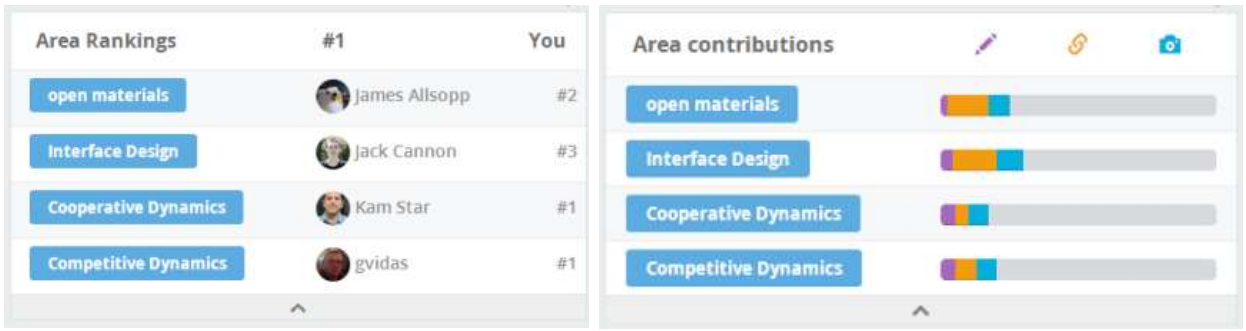


Figure 3.5: *StarQuest's* Leaderboard under Competitive (left) and Cooperative (right) Conditions.

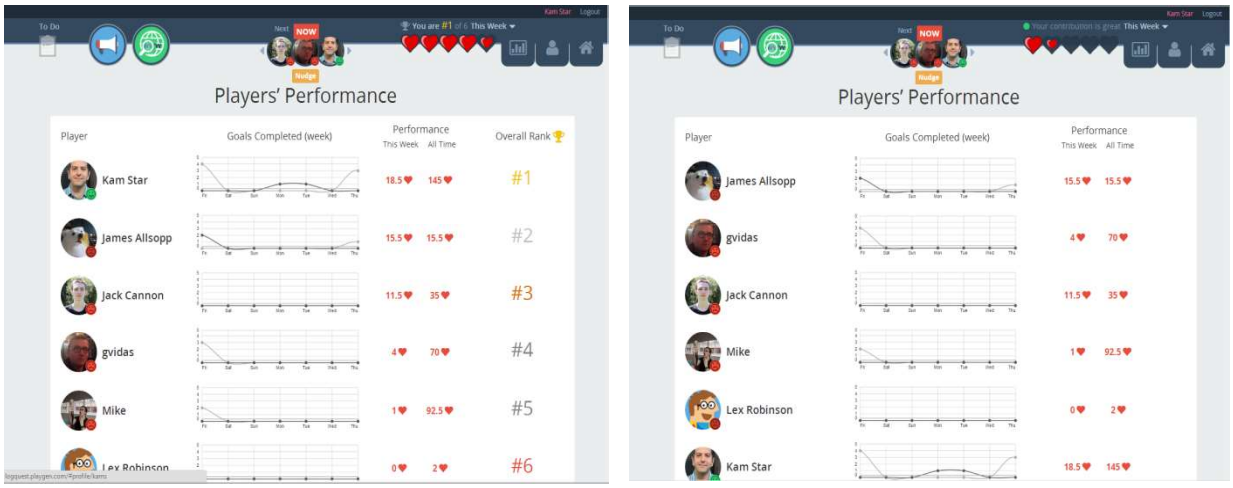


Figure 3.6: *StarQuest's* Performance Page.

Status Mechanism, Smileys as Feedback of Activity or Lack Thereof

The use of virtual human faces in computer interfaces to improve feedback and increase ease of use and communication has been a topic of considerable research (Walker et al. 1994). They have been employed in many areas: online learning (Moreno et al. 2001), advertisements (McBree and Jack 2001), and computer games (Cassell 2004). Called 'embodied agents,' smilies, humanoid synthetic agents, emojis, and emoticons, they have been shown to significantly raise users' positive perceptions of their experience with a system (Mulken et al. 1998). Using 46 studies, Yee et al. (2007) performed a meta-analysis of the impact of their inclusion and their level of realism on user experience. According to the authors' findings, their use increased positive social

interactions with respect to both subjective and behavioural performance measures; level of realism of agents was found to have no impact on behavioural outcomes however (Yee et al. 2007).

StarQuest incorporates human face ideograms in the form of simple 'smileys' as a feedback mechanism to indicate 'happiness' status. The ideogram has five forms, indicating stages ranging from a sad, angry red face to an extremely happy-appearing green face. Depending on interdependent condition, happiness is displayed at group level (cooperative condition) or at individual level (competitive condition). As shown in Table 3.5, the control condition has no happiness status indicator.

Table 3.5: Happiness Status by Condition.

	Cooperation Condition	**Competition Condition**	**Control Condition**
Happiness Status	Status displayed for group as a whole. Group happiness increases with individuals' efforts decreases with their lack of effort.	Status is displayed at individual level. Individual's effort impacts his/her happiness status solely.	No happiness status displayed.

3.2.1.2 Dependent Variables

Previous research on online collaboration platforms differentiates between individuals who add new content, sometimes referred to as 'posters,' and those who read posts but do not add content, referred to as 'lurkers' (Nonnecke and Andrews 2004, Han et al. 2012). Study dependent variables with respect to performance are measured implicitly through participant actions in *StarQuest*. Thus, each time a user visits the StarQuest site, adds a post or a comment, the activity and length of time spent on the platform are stored.

The Big Five Inventory (BFI) measures levels of the five personality traits within the Five Factor Model (i.e., Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism). Developed by John et al. (1991), the BFI instrument contains 44 short phrases which participants rate and can be completed in five to ten minutes. The BFI was selected instead of the short form of NEO-PI-R developed by Costa and McCrae (1992), which contains 60 items, because, while faster to complete, it has been shown to have internal consistency, convergent validity, and discriminant validity (Mullins-Sweatt 2006). Participants were required to complete the BFI before being given access to *StarQuest's* main user interface and after acknowledging informed consent.

3.3 Data Collection, Participants, and Ethics

This study's purpose was to explore the impact of gamification of a social platform, incorporated through reward structures, on social creativity and learning within the context of group collaboration. The platform's design process was human-centric, employing a cohort of end users throughout the process.

3.3.1 Participants

Generalisability of any study's findings relies largely on their applicability under a variety of conditions and in populations besides those employed in the experiment. According to Campbell and Stanley (1963), running an experiment with a variety of participants and under differing conditions increases a study's external validity, the extent to which its findings can be generalised. *StarQuest* was therefore presented to a range of course leaders at Coventry University, where this Ph.D. is being conducted, in order to obtain study participants. From the courses whose leaders were approached, three courses consisting of 363 first- and second-year undergraduate students were invited to participate in the study, which was conducted between September 2014 and November 2014.

3.3.2 Data Collection Methods

The study's experimental data were collected through the participants' use of the *StarQuest* platform. Each user was given a unique login to the system, and all participant actions on the platform were recorded and stored in a database. Additionally the personality trait questionnaire was incorporated into the platform, thus making possible the running of complex queries involving the ranges of the experiment's independent and dependent variables.

The use of website analytics software to measure the usability and performance of online platforms is well documented (Clifton 2012, Fang 2007). According to the most recent global usage statistics, as of August 2014, 49.6% of all online websites employ Google Analytics (Google 2015). Incorporating this analytics software allowed the platform to collect data on length of visit, even if a participant was simply browsing and performed no action (i.e., adding new posts or commenting).

3.3.3 Data Security

Given the nature of the data being collected through the online platform, data privacy and security had to be robust and the communication of any personal information concerning a user to a third party had to be closely scrutinized. Thus, the following precautionary measures were instituted:

- The server and data environments were separated through a separate filesystem for the experiment, and any attempts to access data outside the platform's system was blocked and logged.
- Security consultants SecTheory and Sucuri tested the platform for vulnerabilities, both externally through network connections and internally by scanning for known vectors and exploits.
- Data security was achieved to a level similar to that mandated by ISO/IEC 27001 and ISO/IEC27002, the international standard to preserve system confidentiality, integrity, and availability (ISO 2015).

3.3.4 Ethical Considerations

Appropriate sensitivity towards any study's participants is critically important in order to conduct ethical research. All types of instruments whose purpose is persuasion inherently possess the potential for misuse. In the broad field of persuasive technologies wherein gamification can be situated, concerns over data privacy and power relations (Fogg 2002), trust (Nickel and Spahn 2012), rise of social conflict (Crammer 2011), addiction (Griffiths and Meredith 2009), and unintended bad behaviour (Pavlus 2010) were noted and acknowledged.

Within recent times, a vocal opponent of gamification on ethical grounds is Ian Bogost, himself a game designer, who labelled gamification 'exploitationware' (Bogost 2011). Exploitation, the unfair treatment of another in order to benefit or gain an unfair advantage for oneself, is inherently unethical. Depending on the circumstances, exploitation may be viewed as harmful or mutually advantageous and may be consensual or nonconsensual. Wertheimer and Zwolinski (2013) state that harmful and/or nonconsensual exploitation is clearly unethical. However a transaction between two parties, e.g., a system and its users, that is mutually beneficial is not unethical provided no third party is harmed. In effect, interactions that are free from negative

externalities, consensual, and mutually advantageous are ethically neutral (Wertheimer and Zwolinski 2013).

Berdichevsky and Neuenschwander (1999) developed a set of principles for use by designers of persuasive technologies that focuses on intended outcomes, privacy, motivation, methods, and misinformation. The aim of these principles, given below, is the design of systems which do not exploit users:

- Motivations are not deemed unethical if traditional methods of persuasion are employed.
- The creators must consider and assume responsibility for all reasonably predictable outcomes of a system's use.
- The system's creators must respect the privacy of its users to the same extent they regard their own privacy.
- Relaying of personal information about a user to a third party must be closely scrutinized with respect to violation of privacy.
- The creators should disclose their motivations, methods, and the intended outcomes of the system, except when such disclosure would significantly undermine an otherwise ethical goal.
- The use of misinformation should be avoided in achieving the persuasive end.
- The Golden Rule: The creators should never seek to persuade users of something that they themselves would not wish to be persuaded to.

This study's experimental instrument (i.e., *StarQuest*) has been evaluated against the principles set forth by Bredichevsky and Neuenschwander (1999), and neither the activity nor its intended outcomes appear to be unethical.

The incorporation of such gamification elements as instant rewards, leaderboards, turn taking, and competitive framing of an activity may be seen as potential ethical risks. However, previous research has employed these mechanisms (Versteeg 2013) in online games (Sicart 2009), and gamification platforms designed for education (Cronk 2012) and the workplace (Kumar 2013) have also employed them. Moreover, given the age and maturity of the participants, a critical factor with respect to ethics as discussed by Cohen et al. (2013), the risk of a novel experience which would detrimentally affect

them appears low. Furthermore, through involvement of stakeholders and informed consensus, this risk can be virtually eliminated.

The process of obtaining consent took into account the maturity of the participant and emphasized that participation was voluntary. As part of the process, the nature of the research was fully explained. Thus, the requirements needed to achieve 'informed consent,' as defined by Cohen et al. (2013), were met. Lastly, the board of ethics at Coventry University reviewed the project and found it acceptable.

3.4 Experimental Analysis

The nature of the field experiment, including the presence of a control condition, would seemingly classify it as a randomised control trial. However, given the organisational limitations and the multi-level nature of the courses adopting the experiment's platform, a quasi-experimental cluster design was adopted. Control over assignment of treatment conditions was exercised by randomly assigning groups to one of the three conditions (i.e., competition, cooperation, control).

A critical issue involved in using a quasi-experiment is internal validity, due to the inability to compare the treatment and control groups. To overcome this limitation, in each instance the smaller study groups of four to six individuals were randomly assigned to each of the three conditions. By using existing study groups grouped based on alphabetical listing of names, each cohort acted as a random sample of the overall population, without being grouped based on abilities, sex, or ages. Additionally, the FFM enabled comparison of individuals having similar personality traits with respect to the three conditions, thereby providing a baseline for moderating impact of personality.

3.4.1 Multilevel Model (MLM)

Quantitative analysis operationalises abstract theories into concrete models. Thus, a statistical model is developed to mathematically represent the proposed theoretical relationships between gamification conditions and performance as moderated by personality traits. The relationships actually observed within the sample data are assumed to represent those within the general population (Singer and Willett 2003). The choice of statistical analysis methods is derived based on the research question, experimental design, and data structure (Raudenbush 1988).

The type of contextual analysis used in this thesis draws on similar multilevel analytical methods employed in educational research (Miller and Murdock 2007), social psychology (Hox et al. 2010), and computer supported cooperative work (CSCW) (Dewiyanti et al. 2007, Strijbos et al. 2004), exploring the effect of collective or group characteristics on individual performance.

3.4.2 Power and Sample Size

Power of statistical tests relies on sample size, expected effect size, and level of significance. In multilevel models, each level is associated with a sample size, defined as total number of units observed for that level. In this study, the average number of participants at level 1 (the individual participants in a single study group) is five individuals, with classrooms on average consisting of twenty groups. The calculations of sample sizes in multilevel models are highly dependent on estimating power for fixed and random effects, variance components, and cross-level interactions, all of which are notoriously difficult to estimate *a priori* (Hox et al. 2010, Scherbaum and Ferreter 2009). As cited by Kreft (1996), a rule of thumb suggests that to achieve 30/30 a total sample size of 900 is required, but this was not practical in the current study. Although tables and simple calculations for effect size exist (Cohen 1992; Murphy and Myors 1999), these tables cannot be readily translated to power computations for multilevel models (Scherbaum and Ferreter 2009).

Previous research on moderating effect sizes of FFM traits on performance suggest a relatively small effect size falling between $\beta = .05$ and .29 (Barrick et al. 2001, Hurtz and Donovan 2000).

Using the methodology outlined by Scherbaum and Ferreter (2009) and given the fixed size of level 1 (i.e., the group size based on study groups averaging five students), the maximum power achievable appears to level off at $power <.30$ (i.e., $\alpha > .7$) even for level 2 (number of groups) greater than 50, which is far lower than commonly used (i.e., $power =.95$, $\alpha < .05$). This extremely low power increases the potential for Type II error, i.e., failing to reject a false null hypothesis.

Using the formula described by Snijders and Bosker (1993), standard error is calculated as follows:

$$\boldsymbol{Standard\ Error} \leq \frac{\boldsymbol{Effect\ Size}}{\boldsymbol{Z_{1-\alpha/2}+Z_{1-\beta}}} \qquad (3.1)$$

where $Z_{1-\alpha/2}$ is the z score associated with the chosen level of Type I error for a two-tailed test and $Z_{1-\beta}$ is the z score associated with the desired level of statistical power. Employing α=.7 as per the findings of Scherbaum and Ferreter (2009) and β=.25 as a typical effect size as found by Hurtz and Donovan (2000) and Barrick et al. (2001) with respect to FFM traits and performance meta-analysis yields a standard error ≤ .2.

Using methodology described in detail by Gelman and Hill (2007, p. 448) and results derived in similar research on individual performance in multilevel small groups (Van Mierlo and Kleingeld 2010), which suggests an intra-class correlation of ICC≈.35 for performance effects under cooperative and competitive conditions for groups, and estimating SE(γ)≈.2 with the assumption that $\sigma^2 total$≈1, the following formula can be used to calculate the number of groups (*J*) needed, since we know *m* (average number of members in a group) = 5.

$$\boldsymbol{SE(\gamma) = \sqrt{\frac{\sigma^2 total}{J}}\ (1 + (m - 1)ICC)} \qquad (3.2)$$

where

$$\boldsymbol{ICC = \frac{\sigma^2 \alpha}{\sigma^2 \alpha + \sigma^2 y}} \qquad \textbf{(3.3)}$$

Solving for *J* using the estimates above yields *J*≈60, i.e., 60 groups of 5, or 300 individuals, were required to achieve the proper level of power. As can be seen, as variance increases so does the total number of groups required.

3.4.3 Statistical Analysis

The experiment presented in this thesis is multilevel, or hierarchical, in nature, where a small number of individuals, forming the lowest level of the data hierarchy and referred to as level 1 or the microlevel, are nested into groups, referred to as level 2 or the macrolevel. However, level 1 relationships vary as a function of the level 2 variable, i.e., an individual's performance is partially dependent on the performance of others within his or her group; hence the group context of the individual's performance must be taken into account (Bonito and Hollingshead 1997). Performance within groups is both heterogeneous, as it relates to differences among individuals, and homogeneous, as it relates to the performance of other group members. Therefore, if this data were to be treated as comprised of independent samples, as assumed in standard regression and

ANOVA and ANCOVA analyses, standard reliability estimates such as Cronbach's Alpha (1951) would likely yield inaccurate estimates in response to the scale question (Bonito et al. 2012, Raudenbush 2008).

Additionally, the data employed in this study were gathered from groups whose members have interacted with one another prior to measurement and for whom interaction was related to performance within the experiment. With some exceptions (Kenny et al. 2006) (e.g., team work where team members share a common score), students within a classroom are rarely manipulated with respect to interdependence. However, for the purpose of this experiment and consistent with similar research on small groups (Bonito et al 2012), interdependence with respect to process and outcome were explicit factors, and so interdependence among group members was expected to be an influential factor at group level with respect to performance as well as at the individual level.

Prior to development of the multi-layer analysis (MLA) such as that described above, researchers typically employed conventional regression by aggregating data at a higher level or by applying the higher level data characteristics to individuals (Hox et al. 2010). Both approaches created problems with respect to interpretation of results and inferences made, known respectively as ecological and atomistic fallacies. Specifically, ecological fallacy occurs when data is analysed at one level but conclusions are formulated at another, i.e., interpretation of aggregated data at the individual level. Atomistic fallacy describes making inferences at a higher level based on analysis performed at a lower level (Hox et al. 2010). For example, analysis of deviation of scores can be carried out by subtracting the group mean from individual performance scores, which are then analysed using regression. Since regression assumes independence of observations, employing group means as estimates loses within-group variation, together with the group effect.

In statistical terms, a small number of individuals separated into several groups, as present in this study's experimental design, poses challenges for valid statistical inferences, particularly with respect to drawing inferences about the regression coefficients (i.e., individual and group-specific intercepts and slopes related to personality factors) and when estimating variance and covariance components at both the individual and group levels (Raudenbush 2008). These challenges become apparent in attempts to fit multilevel linear models to data incorporating correlated factors such as FFM traits as measured by the BFI (DeYoung 2006). In this case, sensitivity of fixed

regression coefficients to inferences about variance is higher under individual models than under population-average models (Heagerty and Zeger 2000 in Raudenbush 2008).

In summary, since this study's research questions pertain to the *individual level effects of personality on performance,* which itself is moderated by group level variation in performance, a multilevel linear model was adopted.

3.4.3.1 Population-Average Model, Generalised Estimating Equation (GEE)

As an alternative method for accounting for group-level interactions in multilevel models, which analyses group-to-group variation as well as within-group correlation and provides a model for group-specific regression coefficients as a function of group-level variables plus random variation, a population-average model was also considered for the analysis. A population-average model, also known as a marginal effects model, as, for example, the generalised estimating equation (GEE), provides the ability to express the coefficients of responses as a function of covariates averaged over group-to-group random effects, such as inter-group differences in group-level FFM traits. With continuous dependent variables, such as performance factor measurements in the present experiment, the coefficients obtained through GEE and MLM are mathematically equivalent (Burton et al. 1998, Roux 2003).

3.5 Summary

This chapter has justified the postpositivist research philosophy adopted through the strategic decision to conduct an in-the-field empirical study. The methods used for designing, piloting, and refining the experimental instrument were described, as was the operationalisation of the dependent and independent variables, whose experimental validity was established. Additionally the methods used for data collection, selection of participants and related ethical issues, and the statistical analysis methods to be used on the collected data, together with justifications for their selection, were presented.

The next chapter presents the analysis of the data gathered through the experiment, enabling exploration of the study hypotheses and research questions RQ4, RQ5 and RQ6.

4 RESULTS AND ANALYSIS

This chapter presents the results of the main experimental study whose methodology was presented in the previous chapter. Study results corresponding to each set of hypotheses being tested are presented and discussed in detail in four sections. The first chapter section provides participant and group descriptive statistics for FFM personality traits, including means, standard deviations, measures of skew and kurtosis, and Pearson correlation coefficients. Comparisons with respect to FFM traits between the three courses participating in the study are also presented and analysed. The second section presents performance scores, methods of dealing with outliers, non-parametric statistical tests and methods applied to transform performance scores to fit a normal distribution. In the third section, the results of ANOVA and scatterplots of performance, as measured by posts to the *StarQuest* platform, against FFM personality traits are presented. The fourth and last chapter section presents the multilevel mixed analysis together with the results of applying the generalized estimating equation.

4.1 Participant Characteristics

As shown in Table 4.1, 363 undergraduate participants from three courses were invited to participate in the experiment's platform, *StarQuest*, over an eight-week starting in October 2014. Participation was not mandatory and use of the platform did not count toward their formative assessments. Of the invited participants, 69 individuals either did not consent to use of their data or never logged in. The remaining 294 were divided into 57 groups of between 4 and 8 (median 5) individuals and randomly assigned an experimental condition. In the following discussion and analysis, *Course 1* refers to

first-year undergraduate students from a media and communication course; *Course 2* first-year undergraduate students from a computer science course; and *Course 3* second-year undergraduate students from a sports psychology course.

Table 4.1: Participants and Groups from Three Courses.

	Experimental Condition							
	Cooperative		*Competitive*		*Control*		Logged in and Consented	Did not consent/Did not login
	Individuals	Groups	Indiv.	Groups	Indiv.	Groups		
Course 1	64	13	66	12	59	11	189	12
Course 2	22	4	15	3	17	3	54	13
Course 3	18	4	14	3	19	4	51	44
Total	104	21	95	18	95	18	294	69

4.1.1 Descriptive Statistics

4.1.1.1 Five Factor Model (FFM) Personality Measurement

Personality was assessed using the 44-item Big Five Inventory (BFI) (John et al. 1991) prior to participants' use of the *StarQuest* platform, and Table 4.2 below presents the descriptive statistics related to the FFM traits for the individuals who participated in the experiment. All FFM scores were normally distributed with skewness falling between -.52 and .36 (se =.142) and kurtosis between -.18 and .72 (se =.283).

Table 4.2: Five Factor Model Measurement Descriptive Statistics and Correlations.

Personality Factors	Descriptive Statistics				Pearson Correlations			
	M	SD	Skewness	Kurtosis	1	2	3	4
1. Openness	.641	(.114)	-.110	.488				
2. Conscientiousness	.670	(.142)	.162	-.177	.346**			
3. Extraversion	.624	(.168)	-.165	.016	.379**	.400**		
4. Agreeableness	.739	(.139)	-.516	.726	.273**	.359**	.113	
5. Neuroticism	.393	(.167)	.362	.330	-.230**	-.410**	-.360**	-.360**

N = 294 ; significance * p < .05 ; ** p < .001 (1-tailed)

Reliability testing for internal consistency of trait subscales was performed using Cronbach's alpha. The Openness subscale consisted of 10 items (α=.755); the Conscientiousness subscale of 9 items (α=.813); the Extraversion subscale of 8 items (α=.797); the Agreeableness subscale of 9 items (α=.833); and the Neuroticism subscale of 8 items (α=.882). All factors were within good to acceptable ranges of reliability (George and Mallery, 2003). Shapiro-Wilk tests of normality and Levene's test of homogeneity of variance did not yield statistically significant results, indicating that the FFM trait measurements were homogeneous and normally distributed. Pearson correlations were comparable to results obtained in similar research using the 44-item BFI scale (Herrmann and Pfister 2013, Reid-Seiser and Fritzche 2001).

4.1.1.2 FFM Comparison between Courses, Genders, and Gamification Groups

A one-way between-group ANOVA employed to compare FFM trait scores between *Course 1*, *Course 2*, and *Course 3* participants revealed a statistically significant difference in Openness with F(3,301)=11.184, $p<.0005$. Despite reaching statistical significance, the actual difference in mean scores between the groups was not large. The effect size, calculated using eta squared, was .098, a relatively medium effect (Cohen, 1998, pp. 284). Post-hoc comparisons using the Tukey HSD test indicated that the Openness mean score for *Course 3* (M=.57, SD =.10) was significantly different from those of *Course 1* (M=.66, SD=.12) and *Course 2* (M=.62, SD=.10). There were no other significant differences of participant FFM between *Courses* 1, 2, and 3.

To investigate potential differences *within* genders between courses, a one-way between-group ANOVA for FFM traits of male participants revealed statistically significant differences between them in Openness with F(3,160)=8.892, $p<.0005$ and in Conscientiousness with F(3,160)=3.201, $p=.025$. Effect size for Openness was large (eta squared =.142) and for Conscientiousness was small (eta squared = .045). There were no significant differences for FFM traits between female participants across courses. A post-hoc comparison using the Tukey HSD test revealed that the Openness mean score for males in *Course 3* was lower by -.104 (95% CI, -.152 to -.057, $p<.0005$) than that of *Course 1* and that of *Course 3* was lower by -.063 (95%,-.120 to-.007, $p=.022$) that *Course 2's*.

The relative lowness of the Openness to Experience measurements of *Course 3* male participants, who were from a sports psychology course, compared to that of males from the other courses (students from media and communication and from computer science

classes) has been found in similar research comparing student FFM personality traits where students were from different undergraduate courses (Alexander et al. 2011), suggesting that individuals choosing to study sports psychology may be less open to new ideas and experiences than those studying such topics as media and communication and computer science.

4.1.1.3 FFM Trait Comparison between Gamification Groups

No significance was found in one-way between-group ANOVAs of FFM trait scores between the three experimental conditions. Sub-dividing the participants along gender and re-running the ANOVA also found no significant gender differences between FFM trait scores across conditions.

4.2 Analysis of Raw Performance Scores

The participants collectively posted 2,589 posts, commented 686 times, and spent 38,180 minutes (636 hours) engaging with the platform *StarQuest* over eight weeks. Performance measurements consisted of three quantitative measurements: number of posts, number of comments, and total amount of time spent on the platform. Table 4.3 presents the means, standard deviations, skewness, and kurtosis for these performance counts. Performance scores for *posts* had significantly positive kurtosis, with values ranging from 42.30 to 2.58 (i.e., greater than 1.96). The *posts* data were also all positively skewed, with values ranging from 5.12 to 1.25, suggesting significant deviation from a normal distribution and hinting at the existence of outliers.

Table 4.3: Descriptive Analytics of Performance by Course.

	Course 1 **(n= 189)**				***Course 2*** **(n=54)**				***Course 3*** **(n=51)**				**Total**
Performance	M (sd)	Skew	Kurt.	Sum	M (sd)	Skew	Kurt.	Sum	M (sd)	Skew	Kurt	Sum	Count
Posts	10.8(12.8)	5.11	42.30	2,037	6.9(4.9)	1.38	20.1	420	3.2(1.9)	1.24	2.58	132	2,589
Comments	3.5(5.2)	2.12	4.49	656	.19(.48)	2.66	6.63	10	.39 (.73)	1.88	3	20	686
Time (mins)	159(133)	1.56	2.85	30,185	77 (58)	2.23	6.76	4,167	75(75)	1.55	2.06	3,828	38,180

4.2.1 Outliers

Parametric statistical analysis, particularly regression-based algorithms, are sensitive to the presence of outliers. The performance scores were checked for scores that were relatively very high or very low. A test for statistical outliers using Grubbs modified standardized values identified three outliers in *Course 1*, one in *Course 2* and none in

Course 3. The *Course 1* outliers were *posts*=132 *Z*=9.265 (Critical *Z*=3.55, *p*<.001), *posts*=64 Z=5.73 (Critical *Z*=3.55, *p*<.001), and *posts*=45 *Z*=4.07 (Critical *Z*=3.55, *p*<.001), and the *Course 2* outlier was *posts*=54 *Z*=8.934 (Critical *Z*=3.55, *p*<.001). Since these scores are legitimate unbiased sample data, they were transformed by Winsorising (Hawkins1980) to the closest non-outlier values as follows: *posts*=45 Z=2.97 for *Course 1*, for *Course 2* and *posts*=45 Z=2.77. Subsequent calculation of skewness and kurtosis for *posts* provided some improvement. For *Course 1*, skewness=1.4 and kurtosis=2.05; for *Course 2*, skewness=1.38 and kurtosis=3.85; and for *Course 3*, skewness=1.24 and kurtosis=2.58.

4.2.2 Tests of Normality

Q-Q plots of performance scores indicated a non-normal distribution for *posts*, *comments* and *time*. A Shapiro-Wilk test for normality was then performed on all performance measures, and all were found to be statistically significant with *p* ranging from .034 to .0005. Thus, the null hypothesis that the performance scores were normally distributed was rejected.

4.2.3 Nonparametric Tests

4.2.3.1 Between-Course Comparisons of Performance Scores

Given the non-normality of the performance scores and the high variability between the mean and median for the three *Courses*, Kruskal-Wallis H tests were employed to determine whether significant differences between the *Course 1*, *Course 2*, and *Course 3* groups existed.

Based on visual inspection of boxplots, distribution of performance scores was statistically dissimilar for *Course*-based groups. The distributions of performance scores were found to differ across group to a statistically significant level. With n=294 participating and sub-divided with n=189 comprising *Course 1*, n=54 *Course 2*, and n=51 *Course 3*, the following test statistics for each of the variables were obtained: *posts* χ^2 (2)=88.626 (*p*<.0005), *comments* χ^2 (2)=70.702 (*p*<.0005), and *time* χ^2(2)=70.702 (*p*<.0005).

Subsequently, pairwise comparisons were performed using Dunn's (1964) procedure incorporating the Bonferroni *p*-value correction for multiple comparisons. Values are

mean ranks. This post hoc analysis revealed statistically significant differences in *posts* between *Course 1* (167.87) and *Course 3* (77.34) (p<.0005) and between *Course 2* (142.45) and *Course 3* (p<.0005). No significant differences between *Course 1* and *Course 2* for *posts* were found.

The significant differences between *posts* for *Course 1* and *Course 2*, both of which were comprised of first-year undergraduate students, and for *Course 3,* comprised of second-year undergraduate students, may be explained by differences in experience in using new and so unfamiliar platforms between first- and second-year undergraduates. In an informal meeting prior to beginning the experiment, *Course 3* leader stated *"First year undergraduates are more likely to use a new platform, as everything is new in the first year, by the time they are in the second year most students have developed habits and preferred methods of working and thus engagement with a brand new platform (the experiment) would likely be lower in the second year than first year."*

Pairwise comparisons using the same methodology for *comments* revealed *Course 1* (168.68) to be significantly different from both *Course 2* (93.90) (p<.0005) and *Course 3* (93.44) (p<.0005). No significant differences for *comments* were observed between *Course 2* and *Course 3*.

Considering the lack of scores for *comments* for *Course 2* and *Course 3* (median=0), this factor was employed in analysis of performance scores for only *Course 1* participants.

Pairwise comparison using the same methodology for *time* revealed all three courses to be significantly different from each other. *Course 1* (171.24), *Course 2* (111.56), and *Course 3* (98.45) with p value between p=.007 (for *Course 2*) and p<.0005 (for *Course 3*) for all other pairs. Students from *Course 1* spent significantly more time (159 minutes) compared to *Course 2* (77 minutes) and *Course 3* (75 minutes).

4.3 Analysis of Transformed Performance Scores

This section presents the methodology used for transforming performance scores as well as the analysis of performance under the three experimental conditions—cooperation, competition, and control (i.e., no gamification).

4.3.1 Transforming Performance Scores

The four categories relating to statistical validity and impacting causal inference are conclusion, internal, construct, and external validity, all relying on appropriate application of statistical tools (Cook et al. 1990). Statistical literature has shown that non-normality causes heteroscedasticity, i.e., unequal variability of measurements across the range of a predictor variable so that there are sub-populations of the measurement, each having different variabilities. In regression tests, its presence produces biased results (Schweder and Hjort 2002).

Since this study was concerned with measuring moderating factors, parametric analysis including multilevel regression modelling, and moderated linear regression analysis, exploring the efficacy of transforming performance scores to obtain normalised data was necessary. Several methods were attempted, including the commonly used logarithmic, inverse, and power transformations; however these methods did not yield normalised data as indicated by performance of Shapiro-Wilk tests on the transformed data.

Therefore, alternative transformation methods, including the Box-Cox transformation (Box and Cox 1964 cited in Osborne 2010), a two-step rank and inverse normal distribution function (Templeton 2011), and the Johnson transformation were carried out. Subsequent to testing for efficacy using Shapiro-Wilk test of normality, the two-step method was selected as exhibiting the highest Shapiro-Wilk p values.

This method does not change the order of values, and hence statistical conclusion validity for inferences such as p-value parameters remains valid (Templeton 2011, Zeis et al. 2009). In the first step of the transformation, the fractional rank of each performance score from each *Course* was calculated. The distributions of the probabilities were assessed to be uniform through examination of histograms; thus the second step of the transformation was carried out using the inverse normal distribution

function, with mean 0 and standard deviation 100 (Abramowits and Stegun 1964 cited in Templeton 2011). The Shapiro-Wilk test (p>.05) confirmed normality of *posts* and *time* performance scores for all three *Courses*, with actual p values ranging from *posts* low in *Course 3* (p =.07) to *time* high in *Course 1* (p=1.00).

Transforming performance scores using the two-step method (Templeton 2011) has two advantages. First, scores are standardised, thus enabling causal relationships that transcend *Course* to be recognised and predictor variables on the criterion variable to be calculated (Hunter and Hamilton 2006) with the whole sample. Secondly, potential Type II error in subsequent multilevel linear regression analyses is reduced by increasing the variance in measured scores (MacCallum et al. 2002).

Transformed values are denoted as follows in the remainder of the thesis. The transformed variable *posts* is denoted as *tPosts*, the transformed *comments* is denoted as *tComm*, and the transformed *time* is denoted as *tTime*.

4.3.2 ANOVA Test between Gamification Conditions

One-way ANOVA with post-hoc Tukey HSD tests were conducted with all *Courses* combined to determine if overall performance (i.e., ignoring the five-person study group effects) varied across the different conditions. Performance on posts significantly varied between the control condition and both gamification conditions (i.e., competition and cooperation). Calculations utilised transformed performance post scores as described above, i.e. *tposts*, with mean of 0 and standard deviation of 100. Performance *tPosts* scores increased from control (M= -23.01, SD=100) to cooperative condition (M=13.62, SD=99) to competitive condition (M=18.80, SD=95). In that order, the differences between the three groups were statistically significant with F(2, 291)=5.176 , p=.006. A post hoc test using Tukey HSD identified significantly higher means for *tPosts* in competitive and cooperative conditions compared to the control condition. The mean difference between competitive condition and the control condition was *tPosts*=41.8, se=13.9, CI[8,75] p=.010. Similarly between cooperative and control condition, the mean difference was *tPosts*=36.7, se=13.9, CI[4,69], p=.024.

Transformed performance scores for *time* or *comments* did not vary significantly across the three conditions. However, despite not reaching significance, more time was spent in cooperation (M=11.47) than in competition (M=9.98), and under both conditions participants spent more time than under the control condition (M=-12.75). These

findings are presented in Table 4.4 and Table 4.5 and provide evidence to support Q4 by suggesting gamification significantly improves performance as measured by *posts* irrespective of whether competitive or cooperative interdependence applies, in comparison with no gamification. Additionally whilst the difference did not reach significance, the amount of time spent on *StarQuest* by participants under either gamification condition was greater than under the control condition.

Table 4.4: ANOVA of Performance Scores between the Three Gamification Conditions

Score†	n	Mean	95% CI		df	F	Sig.
			Lower	Upper			
tPosts	294	3.45	-7.98	14.87	2,291	5.176	.006
tComm††	189	5.87	-7.07	18.80	2,186	1.423	.244
tTime	294	3.17	-8.39	14.72	2,291	1.767	.173

† Scores are transformed, sd ≈ 100
†† *tComms* score correlation are for *Course 1* only.

Table 4.5: Post hoc Comparison of Performance Scores between the Three Conditions Using Tukey HSD

				Cooperation : j			Competition : j		
Score	Condition (i)	Mean	SD	M Diff. i-j (se)	Sig.	95%CI (L,U)	M Diff. i-j (se)	Sig.	95%CI (L,U)
tPosts	Coop.	13.62	99						
	Comp.	18.80	95	5.17	.927	-28,38			
	Control	-23.01	100	-36.7*	.024	-69,-4	-41.8**	.010	-75,-8
tComm	Coop.	-2.75	98						
	Comp.	20.91	85	23.67	.293	-13,61			
	Control	-1.62	85	1.12	.997	-37,39	-22.5	.344	-60,15
tTime	Coop.	11.47	110						
	Comp.	9.98	93	-1.48	.994	-35,32			
	Control	-12.75	96	-24.2	.207	-58,9	-22.7	.265	-57,12

† Scores are transformed, sd ≈ 100
†† *tComms* score correlation are for *Course 1* only.
* $p < .05$; ** $p < .001$

4.3.3 Scatterplot of Performance (*tPosts*) against FFM Personality Traits

Figure 4.1 displays scatterplots of transformed, centred *posts* against each of the five personality traits, together with simple regression plot lines of mean *tPosts* against mean personality traits for each of the three conditions (i.e., cooperative, competitive, and

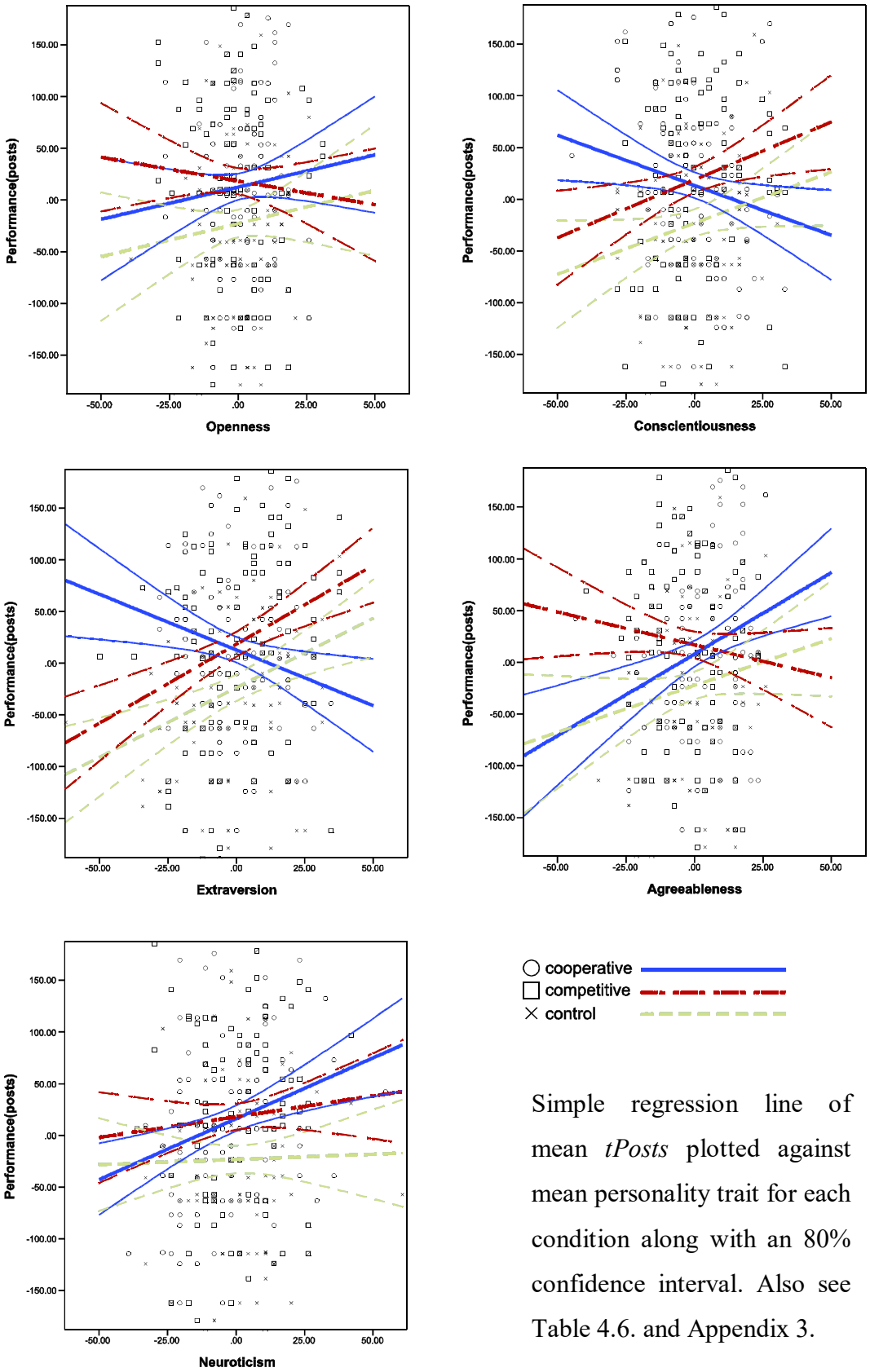
control) each with 80% confidence intervals displayed. As indicated by the plot lines, personality trait appears to impact performance. However, given the wide range of performance scores as measured by *posts*, the relationship does not appear as a well-defined linear association but rather as an overall trend. The regression lines R^2 for all plot lines, representing the goodness-of-fit of the linear models to the data, are presented in Table 4.6. Three values were greater than .05: Extraversion under competition (R^2=.081), Extraversion under Control (R^2=.061), and Agreeableness under cooperation (R^2=.051). Also see Appendix 3 for individual scatterplot charts.

The plots suggest that the relationship between personality and performance is oppositional between cooperation and competition in all traits but Neuroticism. As can be seen in Figure 4.1's panels, under Openness, Conscientiousness, Extraversion, and Agreeableness, the regression lines of cooperation and competition have slopes differing with respect to sign (i.e. as one goes down the other goes up and vice versa). Examining the amount of overlap between the confidence interval plot lines between the five plots suggests that Extraversion is the best candidate as a predictor of performance, at least with respect to the measure *tPosts*, as the clearest differences in performance are exhibited at either end of the personality scale for it.

Table 4.6: Linear R^2 for Mean FFM Trait and Mean tPosts from Scatterplots (see Figure 4.1)

	Cooperation	Competition	Control
Openness	.005	(.003)	.005
Conscientiousness	(.022)	.029	.017
Extraversion	(.025)	.081	.061
Agreeableness	.051	(.009)	.013
Neuroticism	.046	.004	.0002

Values in () brackets indicate a negative slope.



Simple regression line of mean *tPosts* plotted against mean personality trait for each condition along with an 80% confidence interval. Also see Table 4.6. and Appendix 3.

Figure 4.1: Scatterplots of Transformed, Centred Performance Measure *tposts* against the Five Personality Traits.

4.3.4 Correlation between Performance and Personality across Conditions

See Table 4.7 below for the following discussion. Following verification of the non-monotonicity assumption through visual inspection of the scatterplots, Pearson's correlation was calculated between performance measures under each of the three conditions and was positive, highly significant for all, and all either medium or strong (Dancey and Reidy 2004); the minimum correlation, between *tPosts* and *tTime*, was .494 (p<.0005). The correlation between performance scores and FFM personality factors was also calculated. Table 4.7 presents these results also. These correlations, based on the transformed performance variables rather than the originals, mimic the patterns displayed by the originals as displayed in Figure 4.1 and Table 4.6 with respect to sign.

Table 4.7: Performance Measure Means and Correlations; Personality Trait Correlations.

					Personality Factors				
Condition	n	Score †	Mean †	*tTime*	Open.	Cons.	Extra.	Agree.	Neuro.
Cooperative	104	*tPosts*	13.6	.494***	.070	-.147 *	-.159 *	.225 **	.215**
	64	*tComms* ††	-2.75	.497***	.012	-.143	-.230 **	.312 **	.070
	104	*tTime*	11.5		.209**	.134*	-.133*	.336***	-.084
Competitive	95	*tPosts*	18.8	.580 ***	-.059	.169 *	.285 **	-.097	.066
	66	*tComms* ††	20.9	.664 ***	.009	.144	.145	.006	.092
	95	*tTime*	9.99		-.041	.001	-.042	-.032	.216 **
Control	95	*tPosts*	-23.1	.531***	.070	.131*	.246**	.112	.016
	59	*tComms* ††	-1.62	.530 ***	.147	.161	.129	.097	.001
	95	*tTime*	-12.7		-.065	-.028	-.051	.148*	.013

† Scores are transformed. Mean for population is 0, sd ≈ 100
†† *tComms* score correlation are for *Course 1* only.
* p < .1 ; ** p < .05 ; *** p < .001
Open. = Openness, Cons. = Conscientiousness, Extra. = Extraversion, Agree. = Agreeableness, Neuro. = Neuroticism

The results shown in Table 4.7 support Q6 as personality moderates performance. In addition, the direction of correlations for competition and cooperation, as shown in Table 4.7, are oppositional for *tPosts* in all but neuroticism as predicted by H_O, H_C, H_E, and H_A and reflected in Figure 4.1. Thus, for instance, the correlation between Openness and *tPosts* under cooperation is .070, whereas that under competition is -.059. This mimics the corresponding results for the original variable *posts* as shown in Table. 4.6.

As an additional test, correlations between non-transformed raw values of performance scores and personality traits were also calculated using non-parametric Spearman's rank-order (rho). Since the *Courses* had different mean performance scores, the calculations were carried out separately for each *Course* as with the data for all *Courses* combined. The results were very similar (+/- .05) in terms of effect size and significance. These tables can be found in Appendix 3.

4.4 Multilevel Mixed Analysis

To test the hypotheses regarding the moderating effect of individual students' personalities on the relationship between gamification conditions and performance, individual students were viewed as nested within small groups. Since measures associated with individuals in a group are not independent due to shared group effects (Schellens et al. 2005), the assumption of independence required for application of traditional ordinary least squares regression analysis and analysis of variance were violated. As a result, multilevel modelling (MLM) was employed to test the hypothesis, as MLM provides the ability to incorporate the effect of within-group interdependency into the analysis (Hox and Kreft 1994, Raudenbush 2008). Additionally as highlighted in the methodology chapter, due to the high level of correlations between FFM traits as measured by the BFI instrument (DeYoung 2006), a marginal effects GEE was also carried out.

As previously discussed, performance within the study was measured with three observed variables—posts made, comments, and time in the system. However, confirmatory factor analysis (CFA) did not yield a satisfactory single measure based on these. In order to avoid duplicative explanations, only the performance variable *posts* is examined in full, with the results associated with the other performance variables *comments* and *time*, derived using the same methodology, reported in short form.

IBM SPSS version 22 was used for the analysis, and, following established methodology for a multilevel model (Heck et al, 2013), a step-by-step approach was adopted. As recommended by Snijder and Bosker (1999, pp. 56-67), since the transformed performance scores were normally distributed and the number of level 2 (i.e., group) units was relatively small whilst the number of fixed effects (i.e., the five FFM traits and the three gamification conditions) was large, restricted maximum

likelihood estimation (REML) was employed in estimating the multilevel model's parameters.

In multilevel analysis, choice of covariate structure influences the size of effects and their statistical significance. In order to select the appropriate covariance structure, a methodology outlined by Smith (2011) and Gurka (2006) was followed using the goodness-of-fit statistics 2LL, AIC, AICC, and BIC to select which of the several candidate covariance structures best fit the data while keeping the structural variables of the model the same. The unstructured, Toeplitz, variance component, autoregressive, and compound symmetry structures were tested, together with their heterogeneous variations where available. Heterogeneous covariate structures, in particular first-order heterogeneous autoregressive covariance, produced the best fitting model; however due to lack of existing literature on personality traits specifying the use of heterogeneous covariates, unstructured covariate structures were adopted instead. Unstructured covariates appear to be more commonly adopted by researchers in small group analysis (Bonito et al. 2012). Research on the use of multilevel statistical analysis for experiments incorporating personality traits suggests that if significance tests of fixed effects do not vary as a function of inclusion or exclusion of random error terms, the covariance matrices adopted are inevitably having to "play the hand you are dealt" (Nezlek 2011, p. 21), even though the adoption of unstructured covariates resulted in lower statistical significances being derived in models where random effects for FFM traits were considered compared to heterogeneous covariates. In this respect, a future direction of exploration in statistical research may be the development of covariate structures specific to FFM.

Table 4.8 contains parameter estimates for the four models presented and referred to throughout this section.

In developing the multilevel model, Model 1, as shown in Equation (4.1) below, specifies the null, or no predictor, model, which provides an estimated mean performance score for all groups, partitioning the variance in performance into within- and between-group components. Model 1 can be expressed in mixed model form as follows:

$$\boldsymbol{Performance_{ij} = y_{00} + u_{0j} + e_{ij}} \qquad \textbf{(4.1)}$$

The model shows how much group *j's* performance affects individual student *i's* performance, which consists of three components: y_{00} = average group performance or

grand mean, u_{0j} = the extent group *j*'s mean differs from the grand mean, and e_{ij} = the variation in student i's performance in group *j*.

Model 1 also provides a measure of dependence of outcomes within each level 2 unit by way of an intraclass correlation coefficient (ICC), which establishes the degree of variance of outcomes incorporating the group effect as contrasted with no group effect.

Table 4.8: Parameter Estimates under Models 1 through 4.

		Model 1	Model 2	Model 3	Model 4
Fixed components					
Intercept	y_{00}	2.21	-22.78**	-25.375 **	-29.08**
FFM	y_{10}				
Openness (O)	y_{10}O		-.131	-.175	-
Conscientious (C)	y_{10}C		.102	.175	-
Extraversion (E)	y_{10}E		1.273***	1.314 ***	-
Agreeableness (A)	y_{10}A		.941*	.856 *	-
Neuroticism (N)	y_{10}N		1.160**	.955 **	-
GAME Conditions					
Control	y_{20}con		-22.78**	-25.375**	-29.08**
Cooperation	y_{20}coo		36.1***	44.05***	39.21**
Competition	y_{20}com		39.1**	38.81**	39.31**
Mean Group FFM	y_{01}FFM				
mOpenness	y_{01}O			-2.921	-3.5**
mConsciouncious.	y_{01}C			.025	1.046
mExtraversion	y_{01}E			-.321	.034
mAgreeableness	y_{01}A			.899	.273
mNeuroticism	y_{01}N			1.707	1.290
GAME x FFM	y_{21}				
Cooperation x O	y_{21}O*coo				1.841*
Competition x O	y_{21}O*com				-.530
Control x O	y_{21}O*con				.214
Cooperation x C	y_{21}C*coo				-.998
Competition x C	y_{21}C*com				.443
Control x C	y_{21}C*con				.910
Cooperation x E	y_{21}E*coo				-.565
Competition x E	y_{21}E*com				1.976***
Control x E	y_{21}E*con				1.727***
Cooperation x A	y_{21}A*coo				1.926**
Competition x A	y_{21}A*com				.480
Control x A	y_{21}A*con				.036
Cooperation x N	y_{21}N*coo				.922
Competition x N	y_{21}N*com				.858
Control x N	y_{21}N*con				1.494*
Variance of random components					
	τ_0	1983 **	1776**	1622**	-
	σ^2	7897***	7485***	7454 ***	3780**
Model Fit					
Deviance (-2LL)		3512.6	3472	3452.4	3443
ICC (p)		.31			
χ^2 base on -2LL difference with M1			χ^2 (8) = 39.1***	χ^2 (13) = 60.2***	χ^2(38)= 69.6***
Degrees of Freedom		2	10	15	40

* $p < .1$; ** $p < .05$; *** $p < .01$

ICC indicated a significant proportion of variance in performance (i.e., number of posts) between groups as well as significantly varying intercepts across groups, i.e., ICC=.24 or 24% (Wald Z=2.65 p=.004). When considering each of the independent conditions,

the competitive condition's between-group variation was significant at ICC=.36 or 36% (Wald Z = 1.801, *p*=.036). However the cooperation condition's between-group variation was not significant. Since the ICC for competitive condition was 36% and significant and that for both conditions was 24% and not significant, a multilevel model was adopted as the best approach to explain the variability within and between groups. Additionally to test the assumption of difference in outcomes across groups of different sizes, the reliability variable across groups was calculated by factoring in group size. The ICC reliability for the smallest group (i.e., n=3) was 63% and for largest group (i.e., n=9) was 90%. Having established the between-group ICC as significant (ICC> .05) (Hayes 2006), the full multi-level model was developed further.

In Model 2, the individual-level *random intercepts-only model*, grand-mean centred FFM traits, and experimental condition GAME were added as fixed effects. Although FFM is made up of five traits, for the sake of simplicity, the five traits are represented simply as FFM in the equations, however in Table 4.8 the first letter of the personality trait is added to the end of the subscript, i.e., $y_{10}E$ represents y_{10} for Extraversion and so on. The Model 2 performance equation, shown in (4.2) below, is as follows:

$$\boldsymbol{Performance_{ij} = y_{00} + u_{0j} + y_{10}FFM_{ij} + y_{20}GAME_j + e_{ij}} \qquad \textbf{(4.2)}$$

As can be seen, in addition to Model 1's elements, Model 2 contains additional terms. GAME = experimental condition (i.e., cooperation, competition, or control). FFM_{ij} represents cross-level interaction between level 1 and level 2, y_{10} represents the group-level average coefficient with variation fixed to zero at level 2; y_{20} represents variation in intercept for the group average of GAME with variation fixed to zero at level 2; and e_{ij} represents variation in estimating individual performance within groups.

Model 2's -2Log Likelihood (-2LL) deviance has decreased by 39.1 to -2LL=3472.87 from Model 1's. The two models differ by 8 degrees of freedom; Model 1 has 2 degrees of freedom (1 fixed effect intercept and 1 random effect intercept) whereas Model 2 has 10 degrees of freedom (1 fixed effect intercept, 3 fixed effects corresponding to the three GAME conditions, and 5 fixed effects for the FFM traits, as well as 1 random effect intercept). As a result, the difference in deviances is distributed as χ^2 with 8 degrees of freedom under the null hypothesis that the variance of the random intercept component is 0. For $\chi^2(8)=39.1$, the *p* value was less than .001, leading to the rejection of this null hypothesis.

The effect of y_{10}FFM, i.e. group-level average personality coefficient, was significant for Extraversion, Agreeableness, and Neuroticism. For Extraversion, y_{10}E=1.249 (se=.373), t(272.72)=3.411, p =.001; for Agreeableness y_{10}A=.941(se=.44), t(273.96)=2.127, p=.034, and for Neuroticism y_{10}N=1.16 (se=.38), t(278.42)=3.047, p=.003. Thus, the null hypotheses that Extraversion, Agreeableness, or Neuroticism have no effect on the relationship between performance and gamification condition were rejected.

The effect of y_{20}GAME was also significant when comparing the intercept under the neutral (control) condition with y_{20}cont=-22.78 (se=13.41), t(50.796)=-1.69, p=.003; under Cooperation with y_{20}Coop=36.1 (se=18.52), t(53.25)=1.949, p = .003; and under Competition with y_{20}Comp=39.09 (se=19.01), t(50.81)=2.057, p=.001.

In Model 3, the *group-level (level 2) random intercepts model*, the group-level variable to explain variability in intercepts across groups was added, where y_{01} represents variation in FFM group mean. Model 3, (4.3) below, is as follows:

$$\boldsymbol{Performance_{ij} = y_{00} + y_{01}FFM\_mean_j + y_{20}GAME_j + u_{0j} + y_{10}FFM_{ij} + e_{ij}} \qquad \textbf{(4.3)}$$

This model is significantly different from Model 1 as it has 15 degrees of freedom due to the addition of FFM_mean, which represents the means for the five traits. The presence of this term thus adds 5 to the degrees of freedom. $\chi^2(13)=60.2$ $p<.001$ and significantly different from Model 2, $\chi^2(7)=21.1$, $p<.005$. However none of the effects for y_{01}FFM_mean, the FFM group mean, was statistically significant.

Finally, in Model 4, the interaction effects between group and individual level FFM and between GAME and individual level FFM and group-level variance for individual levels of FFM were added as follows in (4.4):

$$\boldsymbol{Performance_{ij} = y_{00} + y_{01}FFM\_mean_j + y_{20}GAME_j + u_{0j} + y_{10}FFM_{ij} + y_{11}FFM\_mean_j * FFM_{ij} + y_{21}GAME_j * FFM_{ij} + u_{1j}FFM_{ij} + e_{ij}} \qquad (4.4)$$

In addition to Model 3's terms, Model 4 incorporates $u_{1j}FFM_{ij}$, which represents the group-level variance for individual level FFM-performance slope; $y_{21}GAME_j * FFM_{ij}$, which represents the slope of interaction effect of GAME conditions with FFM at individual and group levels; and $y_{11}FFM\_mean_j * FFM_{ij}$, which represents the slope for the interaction effect of FFM condition at group with individual level.

The likelihood ratio test for Model 4 compared to Model 1 was significant, with $\chi^2(38)=69.6$, $p<.001$. Following the methodology of Hayes (2006), comparing the Q^2 of

Models 4 and 1 reveals the proportion of the variance remaining after between-group differences are accounted for that can be explained by participant individual personality and by relative personality within the group. This level 1 pseudo R^2=.52, analogous to a squared partial correlation, is the proportion of variance relative to variance in performance remaining after the difference between groups and interactions between game condition and personality are taken into account. In other words, the combination of game condition and personality accounts for 48% of the variation in performance as measured by posts.

Thus, after controlling for an individual's personality interaction with those of the rest of the group and for the interaction between FFM factors and game condition, all five personality factors appear to moderate posts. However only Extraversion and Agreeableness had significant values of p<.05.

As predicted in H_{O1}, Openness positively impacted number of posts under cooperation although not significantly, with y_{21}O*coo=1.841, t(266) = 1.694, p=.091, and, as predicted in H_{O2}, Openness negatively impacted performance as measured by posts under competition although not significantly, with y_{21}O*com=-.532, t(266)=-.61, p=.542. With respect to H_{O3}, performance of individuals with higher Openness scores improved under competitive conditions.

As predicted in H_{C1}, Conscientiousness negatively impacted posts under cooperation with y_{21}C*coo=-.998, t(266)=-1.129, and as per H_{C2}, positively impacted performance under competition with y_{21}C*com=.384, although the interactions were not statistically significant. Conscientiousness also positively impacted the control condition, without reaching statistical significance at p=.05.

As predicted in H_{E2}, Extraversion significantly and positively impacted performance under competition with y_{21}E*com=1.976, t(266)=2.744, p=.006, and, as per H_{E1}, Extraversion negatively impacted performance under cooperation with y_{21}E*coo=-.565, although the interaction did not reach statistical significance at p<.05. Extraversion also positively impacted performance as measured by number of posts under the control condition with y_{21}E*con=1.727, t(266)=2.517, p=.012.

As predicted in H_{A1}, Agreeableness significantly and positively impacted performance under the cooperative condition with y_{21}A*coo=1.926, t(266)=2.309, p=.022. Contrary to the prediction of H_{A2}, Agreeableness positively impacted performance as measured by number of posts under competition with y_{21}A*com=.480, although this result was

not statistically significant. Under the control condition, Agreeableness also positively impacted performance as measured by number of posts with y_{21}A*con=.036, although again the interaction did not reach statistical significance at p=.05.

Neuroticism impacted performance positively under all conditions, with y_{21}N*com=.922 under competition and y_{21}N*coop=.858 under cooperation although these were not significant. Only under the control condition did the interaction approach statistical significance with y_{21}N*con=1.493.

4.4.1 Comparison of Competition and Cooperation as Moderated by Personality Traits

Starting with Model 4 and substituting each of the five factors for FFM and each of the three conditions for GAME, the equation included 50 parameters. Given such a high number and consequently a larger variation than assumed in chapter 3 in estimating sample size, the sample of groups j=57 and participants i= 294 proved insufficient for calculation of all parameters, in particular the random variance coefficients for the personality factors in just one calculation. Additionally, since the previous calculations included the control condition as well as the competitive and cooperative conditions, the estimated variables did not provide a clear comparison between competitive and cooperative conditions. Thus, a more parsimonious approach was adopted, using the same mixed-level methodology as described earlier but reducing the number of variables to include only participant groups under competitive and cooperation conditions and considering only one FFM factor at a time. In this way, model parameters were reduced to 10. The results of this model are shown in Table 4.9 below.

Table 4.9: Model 4 Parameter Estimates after Model Modification to Reduce Number of Parameters.

		Model 4 FFM Personality factors					
		Null Model	1. Open.	2. Cons.	3. Extra.	4. Agree.	5. Neuro.
Fixed components							
Intercept (comp.)	y_{00}	13.6	13.68	13.59	9.62	11.49	10.92
Cooperation	y_{20}		1.32	-3.62	-5.9	-9.89	16.17
Mean FFM (1 to 5)	y_{01}		-2.26	-.74	-1.9	-.241	2.79 **
FFM x M_FFM			-0.47	.035	.093	.1	-.30
Coop. x FFM	y_{21}		1.369	-.921	-.315	1.08	.41
Comp. x FFM	y_{21}		-.029	1.377	2.316 **	.09	-.30
Variance of random components							
	τ_{00}	1583**	1718**	1837*	1789**	1734**	1441*
	τ_{01}		-16.2	14.98 *	41.08 *	67.5*	3.68
	τ_{11}		.153	.663	4.34	2.63	3.14
	σ^2	7832***	7793***	7504**	6432***	7358***	7040***
Model fit							
Deviance (-2LL)		2369	2355	2353	2341	2346	2351

χ^2 based on -2LL difference	$\chi^2(5)=14^*$	$\chi^2(5)=16^{**}$	$\chi^2(5)=28^{***}$	$\chi^2(5)=23^{***}$	$\chi^2(5)=18^{***}$

* $p < .05$ ** $p < .01$ *** $p < .001$

4.5 General Estimating Equation

As outlined in the methodology chapter, a generalized linear model using the general estimating equation was calculated, incorporating information for 199 participants (104 under the cooperative condition and 95 under the competitive condition) within 39 groups and accounting for the interaction between group FFM and individual FFM in the model. Table 4.10 below provides the parameter estimates from GEE. Interaction between gamification condition and Openness as it impacts performance measured by posts was significant with a Wald $\chi^2(1)=4.165$, $p=.041$, where Openness positively impacted cooperation and negatively impacted competition. Similarly the interaction between gamification condition and Extraversion was significant with a Wald $\chi^2(1)=7.806$, $p=.005$, where Extraversion significantly and positively impacted posts under competition and negatively under cooperation.

Table 4.10: Parameter Estimates for *tPosts* under Cooperative and Competitive Conditions Using the Generalized Estimating Equation.

Parameter	β	Std. error	Wald X^2	Sig.
Competitive	6.15	11.35	.281	.596
Cooperative	1.52	15.20	.01	.920
Open. x Cooperative	1.94 *	1.068	3.18	.050
Open. x Competitive	-.773	.854	.82	.365
Cons. x Cooperative	-1.308	.833	2.296	.130
Cons. x Competitive	-.031	.836	.001	.971
Extra. x Cooperative	-.501	.783	.409	.522
Extra. x Competitive	2.159 **	.773	8.66	.002
Agree. x Cooperative	1.420	.828	2.94	.087
Agree. x Competitive	.864	.903	.914	.339
Neuro. x Cooperative	.610	.679	.806	.369
Neuro. x Competitive	.740	.722	1.048	.306

* p < .05 ** p < .01
Open. = Openness, Cons. = Conscientiousness, Extra. = Extraversion, Agree. = Agreeableness, Neuro. = Neuroticism

4.6 Summary

This chapter provided the descriptive and inferential statistics for the empirical study of this thesis. Significant evidence exists to support the inclusion of gamification in *StarQuest* as improving performance. Although the difference between competitive and cooperative conditions was not significant, nevertheless performance as measured by posts was higher under the competitive condition than under the cooperative one, and performance as measured by time was higher under the cooperative condition than under the competitive.

Whilst there is support for the majority of the hypotheses presented in Chapter 2, only Extraversion and Openness appear to be statistically significant in moderating performance as measured by posts between the two interdependent conditions of competition and cooperation. The next chapter discusses these findings in greater detail.

5 CONCLUSIONS

The final chapter of this thesis summarises the main findings of the research and empirical study and reflects on them within the context of the study's aims. The scientific and commercial implications of each finding in the context of the body of work is highlighted. The study's limitations are discussed, and the rich areas of possible research that its conclusions suggest are highlighted.

5.1 Summary of Results

Game elements of points, status, goals and leaderboards were incorporated in the *StarQuest* platform in such a way that activities in the platform were framed as competitive or cooperative interdependent structures (as contrasted with the control condition, which excluded game elements). Performance levels as measured by number of posts, number of comments, and amount of time spent on the platform were captured, stored, and analysed at the conclusion of the study.

The outcome derived from the efforts of the 294 participants in the three courses, who between them spent 636 hours interacting with the three conditions of *StarQuest* (i.e., competition, cooperation and control) provided statistically significant differences in performance as measured by number of posts. Under this performance metric, participants performed 14% better under the cooperative condition than under the control and 16% better under the competitive condition than under the control. These measurements were carried out ignoring the impact of individuals' performance as moderated by their group's performance.

Since the *StarQuest* experiment incorporated the same participant study groupings that the students' tutors had placed them in previously, with a median of five individuals per group, a multilevel analysis was thought to provide a more accurate measure of relative difference in performance. In multilevel analysis, taking into account moderating effects of other participants' performance in the group on an individual's performance, the marginal mean differences were higher and statistically significant at a higher level. In comparing cooperative and control conditions, the mean difference was higher in favour of cooperation with *posts*=40.461, se=15.12, Wald CI[10.8,70.1], *p*=.007, i.e., 15.7% ±6% more posts. In comparing competitive and control conditions, the mean difference was also higher in favour of competition with *posts*=46.696, se=15.249, Wald CI[16.8,76.6], *p*=.002, i.e., 18%±6% more posts. Although the difference between time spent on each of the platform conditions was not statistically significant, results concerning number of posts point to the importance of how time on the platform was spent.

The study results further reinforce suggestions made by other researchers (Hamari et al. 2014) concerning context and perceived utility as a greater determinant of adoption and usage than gamification. Participants were drawn from three undergraduate courses—media and communication (first year), computer science (first year), and sports psychology (second year). In general, first year students tended to use the platform far more than those in second year. The lack of established relationships and familiarity between individuals may account for this, as first year students in their first semester of studies, when the experiment took place, have not typically had as much time as second year students to become acquainted with one another. Another factor determining differences in usage may lie in the contextual use of the platform and its perceived utility by the participants. The students drawn from the media and communications course, by the nature of their interest, tended to be more adept at finding and sharing media on the web, unsurprising since their activities aligned closely with the core functionality of *StarQuest*, i.e., finding and sharing digital media. In comparison to these students, the computer science students spent on average 51% less time and made 33% fewer posts. Also in comparison to the media and communications students, those from the sports psychology course spent on average 50% less time and made 69% fewer posts. In addition, given that computer science and sports psychology students may naturally tend to rely less on sharing rich media such as videos and images than do communication and media students, their perception of a platform dedicated to sharing

such media would represent a lower level of utility than for students from the other courses.

Given the varying efforts that the participants from the different courses expended in *StarQuest*, the findings that gamification improved performance cannot be readily generalised to state that gamification will always improve performance to the same extent. Furthermore, as highlighted previously, one of the critical factors in adoption and use of a platform is its utility in a given context.

5.1.1 On the Role of Social Interdependence, Competition and Cooperation in *StarQuest*

The literature review chapter established the gap in scientific knowledge as it relates to the differences in participant performance engaged in cooperative versus competitive gamification. By using framing techniques in *StarQuest*, each of the gamification conditions was communicated to the participants. Under the competitive condition, participants were encouraged to compete with one another within their group on the basis of individual scores, status, goals, and leaderboards. Under the cooperative condition, presentation of group scores, status, goals, and cooperative leaderboards encouraged them to cooperate in pursuit of group goals and group scores.

Despite not reaching statistical significance, participants in all three courses using *StarQuest* made more posts under the competitive condition than under the cooperative condition. However the difference in performance between cooperative and competitive was very slight, around 1 to 2%. Conversely, participants seemed to spend slightly more time on the platform in the cooperative condition than in the competitive condition (around 1%). Even after counting all 5,895 sessions across the whole experiment, of which 1,441 sessions were under competitive condition and 1,582 were under cooperative condition, the cooperative condition total time was only 10 seconds longer on average than the competitive condition, which lasted on average 6 minutes and 30 seconds. Under the control setting, the average session length was 5 minutes. Even though on the whole performance as measured by posts was slightly higher under the competitive condition as compared to cooperative condition, given the overwhelming evidence in support of using cooperative interdependent structures in education (Johnson 2013), and the evidence provided by *StarQuest* indicating more time being spent on the platform under cooperative condition, the exploration of cooperative framing of game elements in gamification remains advisable.

5.1.2 On the Moderating Effect of Personality Traits on Performance in *StarQuest*

As predicted in the literature review and the subsequent hypothesis formulation, personality traits measured through FFM did indeed appear to moderate individual performance within *StarQuest*, although effect sizes were generally small, a finding that was expected based on the review of related literature, which documented the same result. With respect to traits, significant results were obtained.

Critically the direction of influence for all personality factors with the exception of Neuroticism followed predictions derived from related literature. On the whole, individuals with higher Openness, lower Conscientiousness, lower Extraversion, and higher Agreeableness scores performed relatively better under the cooperative condition than under the competitive condition. Similarly individuals with lower Openness, higher Conscientiousness, higher Extraversion, and lower Agreeableness scores performed relatively better under the competitive condition than under the cooperative one.

The most robust predictors of performance as measured by posts under the two gamification conditions (i.e., competition and cooperation) was Extraversion, where its moderating effect was significant at Wald χ^2 (1)=7.806, p=.005. In all statistical analysis, including Pearson correlations, the multilevel mixed method, and the generalised estimating equations, those higher in Extraversion significantly outperformed those with lower scores in terms of number of posts under the competitive condition, with a small effect size of R^2=.08. Whilst the opposite impact was measured under cooperative condition, with Extraversion reducing the number of posts, the size of the effect was even smaller at R^2=.02 and did not reach statistical significance.

Openness was a significant moderator of performance as measured by posts under both conditions with Wald χ^2(1)=4.165, p=.041. Individuals higher in Openness performed significantly better under cooperation, with a positive effect size of R^2=.07. Although Openness negatively impacted performance, the effect size was even smaller at R^2=.03, and it did not reach statistical significance.

Agreeableness moderated performance positively under both competitive and cooperative conditions. However whilst the interaction was measured as significant under the cooperative condition with effect size of R^2=.05, the effect under competition was less, with R^2=.02, and did not reach significance, thus implying that, whilst

Agreeableness is not a determinant of performance under competitive conditions, it does lead to higher performance under cooperation.

The moderating effects of Conscientiousness and Neuroticism were, however, relatively less significant, whilst the direction of effect for Conscientiousness, following the predictions in literature, did not reach significance. Additionally Neuroticism appears to positively affect performance of posts throughout all conditions.

In addition, parameter estimates from the generalized estimating equation results were used, focusing on Extraversion and Openness only, as only these two traits suggested a significant effect between the conditions and had similar positive and negative impacts in oppositional direction on performance as measured by number of posts. Competition appeared the slightly more promising condition to increase performance overall in *StarQuest*. However if this were to be generally adopted in the educational setting, individuals who are more open, more agreeable, and less extroverted would perform at a relatively lower level. Thus, by shifting competition to between teams and adopting within-team cooperative framings, it may be possible to satisfy the desire for both interdependencies and maximise performance.

The intraclass correlation for the group effect was measured at ICC=.24 or 24% (Wald Z=-2.6, p=.004) indicating that the impact of others within the group was potentially higher than the impact of the individual's personality. This finding implies that the moderating effect of personality on performance is less a predictive measure of performance than the effect composed of the amount of effort others are expending in the group. Thus, other contextual factors, such as efforts of others, may be a higher determinant of performance than personality.

5.2 Study Contribution

As explored in Chapter 2, over 80% of empirical studies in gamification fail to provide significant evidence of improved performance, particularly with respect to the way in which researchers quantify their assertions that "gamification works" (Hamari et al. 2014). Increased motivation has been assumed to increase productivity, all else being equal (Annetta et al. 2010), but the majority of gamification studies do not claim that increased motivation leads to improved performance. Such a discrepancy in reported results within the research literature, further fuelled by claims of gamification proponents, who claim the earning of points can replace the need for valuable interpersonal interaction (Herger 2014, Mollick and Rothbard 2013), create significant problems for the domain over the longer term, as results derived from real world uses fail to live up to claims within the academic literature. Thus, a number of researchers (Bogost 2011, Rey 2015) have strongly objected to the claims of Zichermann and Cunningham (2011), and calls for more nuanced and meaningful approaches (Deterding 2013, Dyer 2015, Nicholson 2015) to incorporating game elements into every-day activities are beginning to be heard.

In addition, recent meta-analyses on the efficacy of serious games to educate concluded that reaching reliable deductions with respect to these claims was impossible (Girard et al. 2012). Similarly, owing to the wildly varied nature of the fields involved in the research, the experimental procedures required, and other contextual elements, there is little scope to prove the effectiveness of gamification in increasing performance in all circumstances.

However, as with some successful serious games, the empirical study reported in this thesis suggests that gamification can succeed in improving outcomes in some situations. While its potential for universal success is far from guaranteed, its findings indicate that, if developed with sensitivity towards end users' genuine needs, gamification can indeed increase performance level. Just as no one game genre suits everyone, no gamification will suit every situation or context, and, as explored in Chapter 2, the core causal links are more likely to be based on the utility of feedback than on extrinsic rewards. Incorporating gaming elements through sensitive and meaningful approaches that focus on adding genuine value to the end-user's experience and provide a vital

function of utility have the potential to increase performance levels, although the sustainability of improved performance is as yet unclear.

Where competition offered marginally better performance with respect to posts, cooperation offered marginally better performance with respect to amount of time spent on the system. Given the evidence in support of individual personality traits leading to preferences with regard to competition and cooperation gamification dynamics, strategies that incorporate personality-based personalised gamification offer the potential of yielding higher performance as compared with the current one-size-fits-all competitive approach adopted by industry. However, given the complexities of developing personalised systems, a general strategy for use of one-size-fits-all approach could involve the incorporation of both types of interdependence, similar to team sports where individuals cooperate with one another whilst competing against other teams, in this way appealing to the broad range of personality traits within individuals which moderate preferences of interdependence structures.

Cooperative interaction remains extremely rare in gamification; as noted earlier some 90% of the empirical studies in the literature review used competition as their primary dynamic for engaging and comparing users. Thus, the development of *StarQuest* provides a methodology for creating cooperative leaderboards and providing group based points, goals, and status. As the use of gamification expands and further matures, and given the evidence provided by *StarQuest* that indicates higher performance by a considerable proportion of individuals based on their personality, it is advisable that designers and researchers of gamification explore the potential of cooperative gamification and make greater use of it in their platforms and experiments.

5.3 Study Limitations

One limitation which possibly reduced the robustness of the study's results was the lack of criticality of outcomes to participants. The *StarQuest* experiment presented a situation of *outcome cooperation* and *outcome competition,* described respectively in detail in Chapter 2 as working independently but winning or losing together and as working independently to compete on outcomes. However, as identified in the methodology chapter, due to ethical concerns and the exploratory nature of *StarQuest*, the outcome or reward from the students' efforts within their courses were not at stake as a result of the experiment. Participation was voluntary, and efforts did not form a basis for any type of summative assessment. Additionally, *StarQuest* did not offer a

method by which means or resources could be made interdependent; participants all had equal access to the platform, thereby enabling them to do as much or as little as any other participant and, whilst the effort was in fact cooperative under all three conditions in that material found by one member of the group was shared with all group members, the only means through which interdependence was manipulated was through rewards. Taken together, these considerations present a challenge to the premise of the experiment with respect to comparing the influence of cooperation and competition. Not only were there no interdependent means, but the interdependence of outcomes was purely within *StarQuest* and thus had no other real-world consequences.

Since the results from the experiment presents *outcome cooperation* against *outcome competition*, as defined in the unified framework for competition and cooperation in Chapter 2, where the outcome in itself is not considered to be critical to the subject, the findings are less likely to be as robust or significant as an experiment comparing *hypercompetition* with *hypercooperation*. This factor may explain the lack of statistical difference in performance between the competitive and cooperative conditions of *StarQuest*.

The multilevel nature of the interactions resulting from participants' study groupings presents another potential limitation of the study. Although effects were measured, the lack of significance in some of the moderating effects of personality traits measured may be due to lack of sufficient number of participants representing each trait under each of the three conditions which would magnify the potential for a Type II. Despite thousands of posts and many hundreds of hours spent on the platform, the relative small number of individuals within each set of interactions meant that, in order to reach statistical significance for some of these effect sizes, larger samples may be required. Nevertheless some significant results were obtained.

5.4 Areas of Future Research

5.4.1 Choice of a Theoretical Framework Underlying Design

The most prominent motivational theoretical frameworks used in gamification concern the delineation between intrinsic and extrinsic motivation followed closely by self-determination theory (SDT) and flow theory. SDT encompasses both intrinsic and extrinsic motivation on a continuum focusing on autonomy, competence and relatedness. However, one promising motivational theory that remains largely

unexplored by gamification researchers is expectancy-value theory, which aims to describe the causal link between action and reward as a driver of behaviour. Supported by empirical studies across the fields of game theory and instructional as well as game design, expectancy-value theory has been considered as a methodology to predict optimum rewards for specific actions in the context of social or individual settings, with further exploration in gamification offering the future potential for developing personalised reward algorithms that respond to an individual's level of effort and expected outcomes as opposed to generic reward systems. The results of this study suggest the addition of personality factors to be considered within this theoretical framework with respect to personality preferences as a predictor of perceived value of an activity within a gamified system.

As indicated earlier in this chapter, goal-setting and goal-setting theory offer one of the most pragmatic frameworks for design of goals within gamification systems as they provide a systematic approach for effectively increase performance. Focusing on pragmatic methodologies such as user-centric design, such as that used in the design process of *StarQuest*, the user-centred theoretical framework for meaningful gamification has been proposed by Nichoson (2012, 2015), offering a theoretical foundation for focusing on end-user needs and goals instead of over the organisation's. Similarly frameworks such as mechanic-dynamic-aesthetic (MDA) are pragmatic tools in gamification design gamification, although as identified in the introductory chapter, the rich tapestry offered by game dynamics such as emotion and narrative remain underutilised in this area.

Recognising the lack of a specific theoretical framework for gamification, a number of authors have proposed new theories as well as user-centred theoretical frameworks for meaningful gamification. Landers and Landers (2014, 2015) proposed a theory of gamified learning, linking specific game elements to learning outcomes and offering empirical methods to test their theory which draws on goal-setting theory and educational research. Finally, by combining intrinsic and extrinsic motivation theories with game dynamics and so-called immersive dynamics, Amir and Ralph (2014) proposed a gamification effectiveness theory, although it requires further empirical validation. In review of gamification-specific theories, Nicholson's (2012) theory of meaningful gamification emphasizes the need for deeply understanding and appreciating the end-user in the design, acknowledging the situational relevance and situated motivational affordance of specific game mechanisms. In the pursuit of

personalised gamification based on personality or individual preferences, this study relied on Nicholson's theory of meaningful gamification as it offers an optimum description through its recognition that game elements and game dynamics are situationally relevant, and, as such, systems that drive them must be designed with sufficient flexibility to be capable of being moulded to the specific user's needs and preferences in order to achieve the greatest effectiveness. As suggested by the results of the *StarQuest* empirical study described in this thesis, and in support of theory of meaningful gamification, personality traits can to some extent predict an individual's affinity toward either a competitive or a cooperative reward structure and as such the potential for further development of similar theories and methodologies of meaningful gamification taking into account personality preferences can be developed toward enhancing gamification's effectiveness in delivering improved outcomes.

5.4.2 Variations on this Study's Interdependence Structures

The unified model of competition and cooperation presented in Chapter 2 offers nine different interdependence situations that ranged from positive interdependence in both means and rewards, termed *hypercooperation*, to negative interdependence of both means and rewards, termed *hypercompetition*, and combinations thereof. The interdependence of goals, resources and rewards presented through the gamification of *StarQuest* provided the opportunity to explore *outcome competition* and *outcome cooperation* as defined in the unified competition and cooperation framework, finding significant results. Thus future gamification research and development may explore the range of other interdependencies offered within the unified framework of competition and cooperation, including *contrariant* and *hyper* variations of competition and cooperation.

Additionally the empirical experiment described in this thesis did not make use of between team interdependence such as inter-team competition or incorporation of inter-team competition together with intra-team cooperation. Given the complexity of personalised systems future experiments may also focus on combination of interdependence by mixing cooperation within teams and competition between teams. Such pragmatic approaches which offer a one-size-fits-all experience are commonplace in such team sports as football and are increasingly being used in entertainment social games such as Clash of Clans to maximise player engagement and performance.

5.4.3 Incorporation of Realistic Risk of Loss and Gain in Experimental Design

A further direction for empirical studies would be to ensure actual interdependence of resources and rewards beyond the reach of an experimental platform, for example by treating the results from *StarQuest* as summative assessment contributing to the students' grades. Although, as noted earlier, such experimentation may run afoul of ethical considerations, as a group of students who receive lower grades as a result of an experimental study where the researcher was aware of the possibility of such effects could be difficult to gain acceptance by a study board.

5.5 Synopsis

This dissertation and the *StarQuest* experiment presented herein point towards the potential for future gamification studies that go beyond the commonly adopted competitive interdependence, by providing evidence in support of both cooperative and competitive elements which appeal to differing personality traits and preferences which moderate performance. Moreover, the presentation of a unified model of cooperation and competition drawn from game theory and social interdependence theory offers the potential for a more nuanced approach to exploration of this domain. Finally this work identifies future directions for gamification research and development, achievable by exploring a range of interdependencies identified in the unified framework of cooperation and competition as well as suggestions of incorporating contrasting interdependencies at both within-group and between-group levels.

6 REFERENCES

Abuhamdeh, S., & Csikszentmihalyi, M. (2009). Intrinsic and extrinsic motivational orientations in the competitive context: An examination of person–situation interactions. Journal of personality, 77(5), 1615-1635.

Ackerman, P. L., Chamorro-Premuzic, T., & Furnham, A. (2011). Trait complexes and academic achievement: Old and new ways of examining personality in educational contexts. British Journal of Educational Psychology, 81(1), 27-40.

Ajzen, I. (1991). The theory of planned behavior. Organizational behavior and human decision processes, 50(2), 179-211.

Alexander, P. M., Pieterse, V., & Lotriet, H. (2011, June). A comparison of computing and non-computing students' personalities based on the five-factor model. In ECIS.

Allport, G. W. (1938). Personality: a psychological interpretation. H. Holt and Company.

Ames, C. (1981). Competitive versus cooperative reward structures: The influence of individual and group performance factors on achievement attributions and affect. American Educational Research Journal, 18(3), 273-287.

Amir, B., & Ralph, P. (2014, May). Proposing a theory of gamification effectiveness. In Companion Proceedings of the 36th International Conference on Software Engineering (pp. 626-627). ACM.

Anderson, J., & Rainie, L. (2012). Gamification and the internet: experts expect game layers to expand in the future, with positive and negative results. GAMES FOR HEALTH: Research, Development, and Clinical Applications, 1(4), 299-302.

Andreasen, M. S., Nielsen, H. V., Schrøder, S. O., & Stage, J. (2007, April). What happened to remote usability testing?: an empirical study of three methods. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 1405-1414). ACM.

Andriole, S. J. (1989). Storyboard prototyping: a new approach to user requirements analysis. QED Information Sciences, Inc.

Annetta L.A., Minogue J., Holmes S.Y. & Cheng M.-T. (2009) Investigating the impact of video games on high school students' engagement and learning about genetics. Computers & Education 53, 74–85

Antonioni, D., & Park, H. (2001). The effects of personality similarity on peer ratings of contextual work behaviours. Personnel Psychology, 54(2), 331-360.

Aronson, E., & Bridgeman, D. (1992). Jigsaw groups and the desegregated classroom: In pursuit of common goals. Readings on the social animal, 430-440.

Atkinson, J. W., & Birch, D. (1970). The dynamics of action.

Atkinson, J. W. (1964). An introduction to motivation.

Axelrod, R. (2006). The evolution of cooperation: revised edition.

Bandura, A. (1999). Social cognitive theory of personality. Handbook of personality: Theory and research, 154-196.

Bandura, A. (1986). Social Foundations of Thought and Action: A Social Cognitive Theory. Prentice Hall, EnglewoodCliffs, NJ.

Barnett, M. A., & Bryan, J. H. (1974). Effects of competition with outcome feedback on children's helping behavior. Developmental Psychology, 10(6), 838.

Barrick, M. R., Mount, M. K., & Strauss, J. P. (1993). Conscientiousness and performance of sales representatives: Test of the mediating effects of goal setting. Journal of Applied Psychology, 78(5), 715.

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next?. International Journal of Selection and assessment, 9(1-2), 9-30.

Barry, B., & Stewart, G. L. (1997). Composition, process, and performance in self-managed groups: The role of personality. Journal of Applied Psychology, 82: 62- 78.

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

Becker-Beck, U., Wintermantel, M., & Borg, A. (2005). Principles of regulating interaction in teams practicing face-to-face communication versus teams practicing computer-mediated communication. Small Group Research, 36(4), 499-536.

Bentea, C. C., & Anghelache, V. (2012). Comparative aspects concerning the effects of extraversion on performance in a cognitive task in competitive and cooperative conditions. Procedia-Social and Behavioral Sciences, 33, 558-562.

Berdichevsky, D., & Neuenschwander, E. (1999). Toward an ethics of persuasive technology. Communications of the ACM, 42(5), 51-58.

Berengueres, J., Alsuwairi, F., Zaki, N., and Ng, T. (2013). Gamification of a recycle bin with emoticons. In Proceedings of the 8th ACM/IEEE International Conference on Human-robot Interaction, pp. 83–84.

Bjork, S., & Holopainen, J. (2005). Games and design patterns. The Game Design Reader, 410-437.

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis.

Blumberg, M., & Pringle, C. D. (1982). The missing opportunity in organizational research: Some implications for a theory of work performance. Academy of management Review, 7(4), 560-569.

Bogost, I. (2015). Why Gamification is Bullshit 2. The Gameful World: Approaches, Issues, Applications, 65.

Bogost, I. (2011). Persuasive Games: Exploitationware. [Online] Available at: http://www.gamasutra.com/view/feature/6366/persuasive_games_exploitationware.php

Bonito, J. A. (1997). Participation in small groups: A performance-based approach (Doctoral dissertation, University of Illinois at Urbana-Champaign).

Bonito, J. A., Ruppel, E. K., & Keyton, J. (2012). Reliability estimates for multilevel designs in group research. Small Group Research, 1046496412437614.

Borman, W. C., & Motowidlo, S. M. (1993). Expanding the criterion domain to include elements of contextual performance. Personnel Selection in Organizations; San Francisco: Jossey-Bass, 71.

Borenstein, M., Cooper, H., Hedges, L., & Valentine, J. (2009). Effect sizes for continuous data. The handbook of research synthesis and meta-analysis, 2, 221-235.

Bouchard, T. J. (1971). Whatever Happened to Brainstorming*. The Journal of Creative Behavior, 5(3), 182-189.

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. Journal of the Royal Statistical Society, B,,26(211-234).

Boyle, G. J. (1995). Myers-Briggs Type Indicator (MBTI): Some Psychometric Limitations. Australian Psychologist, 30(1), 71-74.

Bozarth, J. (2011). An Interview with Sebastian Deterding. eLearn, (7). doi:10.1145/2001333.2008214

Bradley, B. H., Klotz, A. C., Postlethwaite, B. E., & Brown, K. G. (2013). Ready to rumble: How team personality composition and task conflict interact to improve performance. Journal of Applied Psychology, 98, 385-392. doi:10.1037/a0029845

Bradley, B. H., Baur, J. E., Banford, C. G., & Postlethwaite, B. E. (2013). Team players and collective performance: How agreeableness affects team performance over time. Small Group Research, 1046496413507609.

Schreiber, B. B. I., & Brathwaite, B. (2008). Challenges for Game Designers.

Brauner, P., Valdez, A.C., Schroeder, U., and Ziefle, M. (2013). Increase Physical Fitness and Create Health Awareness through Exergames and Gamification. In Human Factors in Computing and Informatics, (Springer), pp. 349–362.

Brewer, R., Anthony, L., Brown, Q., Irwin, G., Nias, J., and Tate, B. (2013). Using gamification to motivate children to complete empirical studies in lab environments. In Proceedings of the 12th International Conference on Interaction Design and Children, pp. 388–391.

Brockner, J., & Higgins, E. T. (2001). Regulatory focus theory: Implications for the study of emotions at work. Organizational behavior and human decision processes, 86(1), 35-66.

Burkey, D. D., Anastasio, M. D. D., & Suresh, A. (2013). Improving Student Attitudes Toward the Capstone Laboratory Course Using Gamification. In American Society for Engineering Education Annual Conf. and Exposition.

Burton, P., Gurrin, L., & Sly, P. (1998). Tutorial in biostatistics. Extending the simple linear regression model to account for correlated responses: an introduction to generalized estimating equations and multi-level mixed modeling. Statistics in medicine, 17, 1261-1291.

Butler, C. (2015). Applied Behavioral Economics: A Game Designer's Perspective. In Gamification in Education and Business (pp. 81-104). Springer International Publishing.

Buunk, B. P., & Brenninkmeijer, V. (2001). When individuals dislike exposure to an actively coping role model: Mood change as related to depression and social comparison orientation. European Journal of Social Psychology, 31(5), 537-548.

Buunk, B. P., Nauta, A., & Molleman, E. (2005). In search of the true group animal: The effects of affiliation orientation and social comparison orientation upon group satisfaction. European Journal of personality, 19(1), 69-81.

Cambria, E., Rajagopal, D., Kwok, K., & Sepulveda, J. (2015, July). GECKA: Game Engine for Commonsense Knowledge Acquisition. In The Twenty-Eighth International Flairs Conference.

Caillois, R., & Barash, M. (1961). Man, play, and games. University of Illinois Press.

Cassell, J. (2004). Towards a model of technology and literacy development: Story listening systems. Journal of Applied Developmental Psychology, 25(1), 75-105.

Cassidy, T., & Lynn, R. (1989). A multifactorial approach to achievement motivation: The development of a comprehensive measure. Journal of Occupational Psychology, 62(4), 301- 312.

Chamorro-Premuzic, T., & Furnham, A. (2003). Personality predicts academic performance: Evidence from two longitudinal university samples. Journal of Research in Personality, 37(4), 319-338.

Charlesworth, W. R. (1996). Co-operation and competition: Contributions to an evolutionary and developmental model. International Journal of Behavioral Development, 19(1), 25-38.

Chen, H. (2006). Flow on the net–detecting Web users' positive affects and their flow states. Computers in Human Behavior, 22(2), pp. 221 – 233. doi:10.1016/j.chb.2004.07.001

Chung, J. E., Park, N., Wang, H., Fulk, J., & McLaughlin, M. (2010). Age differences in perceptions of online community participation among non-users: An extension of the Technology Acceptance Model. Computers in Human Behavior, 26(6), 1674-1684.

Cialdini, R. B. (1993). Influence: The psychology of persuasion. New York: Morrow.

Clayton, K., Blumberg, F., & Auld, D. P. (2010). The relationship between motivation, learning strategies and choice of environment whether traditional or including an online component. British Journal of Educational Technology, 41(3), 349-364.

Clifton, B. (2012). Advanced web metrics with Google Analytics. John Wiley & Sons.

Cochoy, F. (2014). Consumers at work, or curiosity at play? Revisiting the prosumption/value cocreation debate with smartphones and two-dimensional bar codes. Marketing Theory, 1470593114540676.

Codish, D., & Ravid, G. (2014). Personality based gamification-Educational gamification for extroverts and introverts. In Proc. 9 th Chais Conf. for the Study of Innovation and Learning Technologies: Learning in the Technological Era.

Cogliser, C. C., Gardner, W. L., Gavin, M. B., & Broberg, J. C. (2012). Big Five Personality Factors and Leader Emergence in Virtual Teams Relationships With Team Trustworthiness, Member Performance Contributions, and Team Performance. Group & Organization Management, 37(6), 752-784.

Cohen, J. (1992). A power primer. Psychological bulletin, 112(1), 155.

Cohen, L., Manion, L., & Morrison, K. (2013). Research methods in education. Routledge.

Colbert, A. E., Mount, M. K., Harter, J. K., Witt, L. A., & Barrick, M. R. (2004). Interactive effects of personality and perceptions of the work situation on workplace deviance. Journal of Applied Psychology, 89(4), 599.

Collins, R. L. (1996). For better or worse: The impact of upward social comparison on self-evalua- tion. Psychological Bulletin, 119, 51–69.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. Knowing, learning, and instruction: Essays in honor of Robert Glaser, 18, 32-42.

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

Conover, W. J. (1999). Practical nonparametric statistics (3rd ed.). Hoboken, NJ: Wiley & Sons.

Cook, D. J., Mulrow, C. D., & Haynes, R. B. (1997). Systematic reviews: synthesis of best evidence for clinical decisions. Annals of internal medicine, 126(5), 376-380.

Cook, T. D., Campbell, D. T., & Peracchio, L. (1990). Quasi experimentation. In, MD Dunnette & LM Hough,(Eds.) Handbook of industrial and organizational psychology: 491–576. Alto.

Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Popović, Z. (2010). Predicting protein structures with a multiplayer online game. Nature, 466(7307), 756-760.

Chang, A. Y. (2012). Back to the virtual farm: Gleaning the agriculture-management game. Interdisciplinary Studies in Literature and Environment, 19(2), 237-252.

Chen, X. P., Xie, X., & Chang, S. (2011). Cooperative and competitive orientation among Chinese people: Scale development and validation. Management and Organization Review, 7(2), 353-379.

Cornelissen, F., Neerincx, M. A., Smets, N. J. J. M., Breebaart, L., Dujardin, P., & Wolff, M. (2012). Gamification for astronaut training. SpaceOps, Stockholm.

Costa, P., McCrae, R., & Holland, J. (1984). Personality and vocational interests in an adult sample. Journal of Applied Psychology 69, 390-400.

Costa, P. T., & McCrae, R. R. (2008). The revised neo personality inventory (neo-pi-r). The SAGE handbook of personality theory and assessment, 2, 179-198.

Zeynep, A., Cramer, H., Holmquist, L. E., & Rost, M. (2011). Gamification and location-sharing: Some emerging social conflicts.

Creswell, J. W. (2003). Research design. Qualitative, quantitative and mixed methods approaches. 2nd edition. EUA Sage, 196.

Cronk, M. (2012, June). Using gamification to increase student engagement and participation in class discussion. In World conference on educational multimedia, hypermedia and telecommunications (Vol. 2012, No. 1, pp. 311-315).

Crumlish, C., & Malone, E. (2009). Designing social interfaces: Principles, patterns, and practices for improving the user experience. " O'Reilly Media, Inc.".

Csikszentmihalyi, M. (1975). Play and Intrinsic Rewards. Journal of Humanistic Psychology, 15(3), pp. 41 – 63. doi:10.1177/002216787501500306

Csikszentmihalyi, M. (1991). Flow: The psychology of optimal experience. HarperCollins.

Csikszentmihalyi, M. (1997). Flow and the Psychology of Discovery and Invention. HarperPerennial, New York, 39.

Dancey, C., & Reidy, J. (2004). Statistics without Maths for Psychology: using SPSS for Windows, London: Prentice Hall.

Danelli, F. (2015). Implementing Game Design in Gamification. In Gamification in Education and Business (pp. 67-79). Springer International Publishing.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS quarterly, 319-340.

Dawkins, R. (1986). The blind watchmaker: Why the evidence of evolution reveals a universe without design. WW Norton & Company.

Deci, E. L. (1980). The psychology of self-determination. Free Press.

Demerouti, E. (2006). Job characteristics, flow, and performance: the moderating role of conscientiousness. Journal of occupational health psychology, 11(3), 266.

Depura, K., and Garg, M. (2012). Application of Online Gamification to New Hire Onboarding. In Services in Emerging Markets (ICSEM), 2012 Third International Conference On, pp. 153–156.

Derryberry, D., & Reed, M. A. (1998). Anxiety and attentional focusing: trait, state and hemispheric influences. Personality and Individual Differences, 25, 745–761.

Deterding, S. (2014). Eudaimonic design, or: Six invitations to rethink gamification. [online] available from <http://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2466374>

Deterding, S. (2015). The Ambiguity of Games: Histories and Rhetorics of a Gameful World. In Walz, S. P., & Deterding, S. (2015). The Gameful World: Approaches, Issues, Applications. Mit Press.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011, September). From game design elements to gamefulness: defining gamification. In Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments (pp. 9-15). ACM.

Deubel, P. (2003). An investigation of behaviorist and cognitive approaches to instructional multimedia design. Journal of educational multimedia and hypermedia, 12(1), 63-90.

Deutsch, M. (2006). Cooperation and competition. In M. Deutsch, P. T. Coleman, & E. C. Marcus (Eds.), The Handbook of Conflict Resolution: Theory and practice (23–42). San Francisco: Jossey-Bass.

Deutsch, M., & Krauss, R. M. (1962). Studies of interpersonal bargaining. Journal of Conflict Resolution, 52-76.

Deutsch, Morton (1949). A Theory of Co-operation and Competition.

Dewiyanti, S., Brand-Gruwel, S., Jochems, W., & Broers, N. J. (2007). Students' experiences with collaborative learning in asynchronous computer-supported collaborative learning environments. Computers in Human Behavior, 23(1), 496-514.

DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. Journal of personality and social psychology, 91(6), 1138.

Digman, J. M. (1990). Personality Structure: Emergence of the Five Factor Model, Annual Review of Psychology, 41, pp. 417-440.

Drenner, S., Sen, S., & Terveen, L. (2008). Crafting the initial user experience to achieve community goals. In Proceedings of the 2008 ACM conference on Recommender systems (pp. 187–194). ACM.

Driskell, J. E., Goodwin, G. F., Salas, E., & O'Shea, P. G. (2006). What makes a good team player? Personality and team effectiveness. Group Dynamics: Theory, Research, and Practice, 10(4), 249.

Duarte, D.M.C. (2012). Using Visualization Techniques and Gamification to Involve Users in Requirements Elicitation.

Dubois, D.J., and Tamburrelli, G. (2013). Understanding Gamification Mechanisms for Software Development.

Duffy, F. M. (2003). I think, therefore I am resistant to change. Journal of Staff Development, 24(1), 30-37.

Dumitrache, A., Aroyo, L., Welty, C., Sips, R. J., & Levas, A. (2013). "Dr. Detective": combining gamification techniques and crowdsourcing to create a gold standard in medical text. In 12th International Semantic Web Conference Proceedings of the 1st International Workshop on Crowdsourcing the Semantic Web (CrowdSem 2013).

Dunn, O. J. (1964). Multiple comparisons using rank sums. Technometrics, 6, 241-252.

Durlauf, S. N., & Young, H. P. (2004). Social dynamics (Vol. 4). Mit Press.

Dyer, R. (2015). A Conceptual Framework for Gamification Measurement. In Gamification in Education and Business (pp. 47-66). Springer International Publishing.

Earls, M. (2015). Copy, Copy, Copy: How to Do Smarter Marketing by Using Other People's Ideas. John Wiley & Sons.

Earls, M. (2009). Herd: how to change mass behaviour by harnessing our true nature. John Wiley & Sons.

Easterbrook, S. (Ed.). (2012). CSCW: Cooperation or conflict?. Springer Science & Business Media.

Egenfeldt-Nielsen, S. (2005). Beyond edutainment: Exploring the educational potential of computer games. Lulu. com.

Eickhoff, C., Harris, C. G., de Vries, A. P., & Srinivasan, P. (2012, August). Quality through flow and immersion: gamifying crowdsourced relevance assessments. In Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval (pp. 871-880). ACM.

Ekehammar, B., & Akrami, N. (2003). The relation between personality and preju- dice: A variable- and person-centered approach. European Journal of Personality, 17, 449-464. doi:10.1002/per.494

Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. Journal of personality and social psychology, 72(1), 218.

Ellis, A. P. J., Hollenbeck, J. R., Ilgen, D. R., Porter, C. O. L. H., West, B. J., & Moon, H. (2003). Team learning: Collectively connecting the dots. Journal of Applied Psychology, 88, 821-835. doi:10.1037/0021-9010.88.5.821

Eyster, E., & Rabin, M. (2005). Cursed equilibrium. Econometrica, 73(5), 1623-1672.

Falkner, N. J., & Falkner, K. E. (2014, November). Whither, badges? or wither, badges!: a metastudy of badges in computer science education to clarify effects, significance and influence. In Proceedings of the 14th Koli Calling International Conference on Computing Education Research (pp. 127-135). ACM.

Fang, W. (2007). Using google analytics for improving library website content and design: a case study.

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

Felder, R. M., Felder, G. N., & Dietz, E. J. (2002). The effects of personality type on engineering student performance and attitudes. Journal of Engineering Education, 91(1), 3-17.

Ferro, L. S., Walz, S. P., & Greuter, S. (2013, September). Towards personalised, gamified systems: an investigation into game design, personality and player typologies. In Proceedings of The 9th Australasian Conference on Interactive Entertainment: Matters of Life and Death (p. 7). ACM.

Festinger, L. (1954). A theory of social comparison processes. Human Relations, 7(2), 117–140.

Festinger, L. (1962). A theory of cognitive dissonance (Vol. 2). Stanford university press.

Field, Andy (2009). Discovering statistics using SPSS (3rd ed. ed.). Los Angeles [i.e. Thousand Oaks, Calif.]: SAGE Publications. p. 143. ISBN 978-1-84787-906-6.

Frith, J. (2013). Turning life into a game: Foursquare, gamification, and personal mobility. Mobile Media & Communication, 1(2), 248-262.

Fischer, G., Giaccardi, E., Eden, H., Sugimoto, M., & Ye, Y. (2005). Beyond binary choices: Integrating individual and social creativity. International Journal of Human-Computer Studies, 63(4), 482-512.

Fitz-Walter, Z., and Tjondronegoro, D.W. (2011). Exploring the opportunities and challenges of using mobile sensing for gamification and achievements. In UbiComp 11: Proceedings of the 2011 ACM Conference on Ubiquitous Computing, pp. 1–5.

Fitz-Walter, Z., Tjondronegoro, D., & Wyeth, P. (2011, November). Orientation passport: using gamification to engage university students. In Proceedings of the 23rd Australian Computer-Human Interaction Conference (pp. 122-125). ACM.

Fogg, B. J. (2002). Persuasive technology: using computers to change what we think and do. Ubiquity, 2002(December), 5.

Fogg, B. J., Marshall, J., Kameda, T., Solomon, J., Rangnekar, A., Boyd, J., & Brown, B. (2001, March). Web credibility research: a method for online experiments and early study results. In CHI'01 extended abstracts on Human factors in computing systems (pp. 295-296). ACM.

Foster, J. A., Sheridan, P. K., Irish, R., & Frost, G. S. (2012). Gamification as a strategy for promoting deeper investigation in a reverse engineering activity. In American Society for Engineering Education. American Society for Engineering Education.

Frang, K., & Mellstrand, R. (2012). Enterprise gamification of the employee development process at an Infocom Consultancy Company. In Congresso Ibero-Americano em Engenharia de Software.

Frith, J. (2013). Turning life into a game: Foursquare, gamification, and personal mobility. Mobile Media & Communication 1, 248–262.

Furnham, A., Jackson, C. J., & Miller, T. (1999). Personality, learning style and work performance. Personality and Individual Differences, 27(6), 1113-1122.

Gagné, M., & Deci, E. L. (2005). Self-determination theory and work motivation. Journal of Organizational behavior, 26(4), 331-362.

Gartner (2013). Technology Research | Gartner Inc. [WWW Document] 80 Percent of Current Gamified Applications 2013. 16 Aug. 2013 <http://www.gartner.com/newsroom/id/2251015>

Gartner (2015). Technology Research | Gartner Inc. [WWW Document], n.d. URL http://www.gartner.com/technology/home.jsp (accessed 9.25.15).

Gellatly, I. R., & Irving, P. G. (2001). Personality, autonomy, and contextual performance of managers. Human Performance, 14(3), 231-245.

Gelman, A., & Hill, J. (2006). Data analysis using regression and multilevel/hierarchical models. Cambridge University Press.

George, D., & Mallery, P. (2003). SPSS for Windows step by step: A simple guide and

George, J. M., & Zhou, J. (2001). When openness to experience and conscientiousness are related to creative behavior: an interactional approach. Journal of applied psychology, 86(3), 513.

Gibson, J. J. (1977). The Theory of Affordances. InPerceiving, Acting, and Knowing, edited by Robert Shaw and John Bransford.

Gibson, D., & Jakl, P. (2015). Theoretical Considerations for Game-Based e-Learning Analytics. In Gamification in Education and Business (pp. 403-416). Springer International Publishing.

Gnauk, B., Dannecker, L., and Hahmann, M. (2012). Leveraging gamification in demand dispatch systems. In Proceedings of the 2012 Joint EDBT/ICDT Workshops, pp. 103–110.

Goel, L., Johnson, N., Junglas, I., & Ives, B. (2010). Situated learning: Conceptualization and measurement. Decision Sciences Journal of Innovative Education, 8(1), 215-240.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. American psychologist, 48(1), 26.

Goldratt, E. M., Cox, J., & Whitford, D. (1992). The goal: a process of ongoing improvement (Vol. 2). Great Barrington, MA: North River Press.

Google (2015). http://w3techs.com/technologies/details/ta-googleanalytics/all/all

Graepel, T., Herbrich, R., & Gold, J. (2004). Learning to fight. In Proceedings of the International Conference on Computer Games: Artificial Intelligence, Design and Education (pp. 193–200).

Grant, L. A. (2008). Culture change in a for-profit nursing home chain: An evaluation. New York: Commonwealth Fund. Retrieved January 14, 2014, from http://www.commonwealthfund.org/usr_doc/Grant_culturechangefor-profinursinghome_1099.pdf

Gray, E. K., & Watson, D. (2002). General and specific traits of personality and their relation to sleep and academic performance. Journal of Personality, 70(2), 177-206.

Graziano, W. G., Hair, E. C., & Finch, J. F. (1997). Competitiveness mediates the link between personality and group performance. Journal of personality and social psychology, 73(6), 1394.

Griffiths, M. D., & Meredith, A. (2009). Videogame addiction and its treatment.Journal of Contemporary Psychotherapy, 39(4), 247-253.

Grinnell, R. M., Unrau, Y. A., & Williams, M. (2008). Group-level designs.Social work research and evaluation: Foundations of evidence-based practice, 177-204.

Groh, F. (2012). Gamification: State of the art definition and utilization. Institute of Media Informatics Ulm University, 39.

Grubbs, F. E. (1969). Procedures for detecting outlying observations in samples. Technometrics, 11(1), 1-21.

Gurka, M. J. (2006). Selecting the best linear mixed model under REML. The American Statistician, 60(1), 19-26.

Hahn, J. (2012). Mastering software: gamification of a help system (Doctoral dissertation, Auckland University of Technology).

Hair, J. F., Tatham, R. L., Anderson, R. E., & Black, W. (2006). Multivariate data analysis (Vol. 6). Upper Saddle River, NJ: Pearson Prentice Hall.

Hamari, J. (2013). Transforming Homo Economicus into Homo Ludens: a field experiment on gamification in a utilitarian peer-to-peer trading service. Electronic Commerce Research and Applications.

Hamari, J., and Koivisto, J. (2013). Social motivations to use gamification: An empirical study of gamifying exercise. In Proceedings of the 21 St European Conference in Information Systems. Utrecht, Netherlands,.

Hamari, J., Koivisto, J., & Sarsa, H. (2014, January). Does gamification work?--a literature review of empirical studies on gamification. In System Sciences (HICSS), 2014 47th Hawaii International Conference on (pp. 3025-3034). IEEE.

Han, J. Y., Kim, J. H., Yoon, H. J., Shim, M., McTavish, F. M., & Gustafson, D. H. (2012). Social and psychological determinants of levels of engagement with an online breast cancer support group: posters, lurkers, and nonusers. Journal of health communication, 17(3), 356-371.

Harrison, D. A., Mohammed, S., McGrath, J. E., Florey, A. T., & Vanderstoep, S. W. (2003). Time matters in team performance: Effects of member familiarity, entrainment, and task discontinuity on speed and quality. Personnel Psychology, 56(3), 633-669.

Hart, J. W., Stasson, M. F., Mahoney, J. M., & Story, P. (2007). The Big Five and Achievement Motivation: Exploring the Relationship Between Personality and a Two-Factor Model of Motivation. Individual Differences Research, 5(4).

Harvey, O. J., White, B. J., Hood, W. R., & Sherif, C. W. (1961). Intergroup conflict and cooperation: The Robbers Cave experiment (Vol. 10). Norman, OK: University Book Exchange.

Hayes, A. F. (2006). A primer on multilevel modeling. Human Communication Research, 32(4), 385-410.

Hawkins, D. M. (1980). Identification of outliers. London: Chapman & Hall.

Heagerty, P. J., & Zeger, S. L. (2000). Marginalized multilevel models and likelihood inference (with comments and a rejoinder by the authors). Statistical Science, 15(1), 1-26.

Heck, R. H., Thomas, S. L., & Tabata, L. N. (2013). Multilevel and longitudinal modeling with IBM SPSS. Routledge.

Heinzen, T. E., Landrum, R. E., Gurung, R. A., & Dunn, D. S. (2015). Game-based assessment: The mash-up we've been waiting for. In Gamification in Education and Business (pp. 201-217). Springer International Publishing.

Herger, M. (2014). Enterprise Gamification: Engaging People by Letting Them Have Fun: Book 1-the Basics. CreateSpace Independent Publishing Platform.

Herrmann, A., & Pfister, H. R. (2013). Simple measures and complex structures: Is it worth employing a more complex model of personality in Big Five inventories?. Journal of Research in Personality, 47(5), 599-608.

Herzig, P., Strahringer, S., and Ameling, M. (2012). Gamification of ERP Systems-Exploring Gamification Effects on User Acceptance Constructs. Multikonferenz Wirtschaftsinformatik MKWI 793–804.

Herzig, P., Ameling, M., & Schill, A. (2012, August). A generic platform for enterprise gamification. In Software Architecture (WICSA) and European Conference on Software

Architecture (ECSA), 2012 Joint Working IEEE/IFIP Conference on (pp. 219-223). IEEE.

Herzig, P., Ameling, M., Wolf, B., & Schill, A. (2015). Implementing gamification: Requirements and gamification platforms. In Gamification in Education and Business (pp. 431-450). Springer International Publishing.

Hilk, C. L. (2013). Effects of cooperative, competitive, and individualistic learning structures on college student achievement and peer relationships: a series of meta-analyses (Doctoral dissertation, UNIVERSITY OF MINNESOTA).

Hilton, M. (2013). Improving WebIDE through Delightful Design and Gamification.

Holland, J. G., & Skinner, B. F. (1961). The analysis of behavior: A program for self-instruction.

Hollenbeck, J. R., Williams, C. R., & Klein, H. J. (1989). An empirical examination of the antecedents of commitment to difficult goals. Journal of Applied Psychology, 74(1), 18.

Holtzblatt, K., Wendell, J. B., & Wood, S. (2004). Rapid contextual design: a how-to guide to key techniques for user-centered design. Elsevier.

Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. Experimental Economics, 14(3), 399-425.

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2010). Multilevel analysis: Techniques and applications. Routledge.

Hox, J. J., & Kreft, I. G. (1994). Multilevel analysis methods. Sociological Methods & Research, 22(3), 283-299.

Hough, L. M. (1992). The “Big Five” personality variables–construct con- fusion: Description versus prediction. Human Performance, 5, 139–155.

Hsu, S.H., Chang, J.-W., and Lee, C.-C. (2013). Designing Attractive Gamification Features for Collaborative Storytelling Websites. Cyberpsychology, Behavior, and Social Networking.

Hulsebosch, M. (2013). Targeting Gamification Applications to Increase User Participation.

Hunicke, R., LeBlanc, M., & Zubek, R. (2004, July). MDA: A formal approach to game design and game research. In Proceedings of the AAAI Workshop on Challenges in Game AI (Vol. 4).

Huotari, K., & Hamari, J. (2012, October). Defining gamification: a service marketing perspective. In Proceeding of the 16th International Academic MindTrek Conference (pp. 17-22). ACM.

Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. Journal of Applied Psychology, 85(6), 869–879. doi:10.1037//0021-9010.85.6.869

IDC (2015). Social collaboration market set to boom, IDC predicts | Information Age. (n.d.). Retrieved September 22, 2015, from http://www.information-age.com/technology/applications-and-development/2110218/social-collaboration-market-set-to-boom-idc-predicts

Inkpen, K., McGrenere, J., Booth, K. S., & Klawe, M. (1997). Turn-taking protocols for mouse-driven collaborative environments. In Graphics Interface(Vol. 97, pp. 138-145).

ISO 9241-210:2010 - Ergonomics of human-system interaction -- Part 210: Human-centred design for interactive systems. (n.d.). Retrieved July 29, 2014, from http://www.iso.org/iso/catalogue_detail.htm?csnumber=52075

ISO (2015) see http://www.iso.org/iso/iso27001

Jacobs, A., Timmermans, A., Michielsen, M., Vander Plaetse, M., and Markopoulos, P. (2013). CONTRAST: gamification of arm-hand training for stroke survivors. In CHI'13 Extended Abstracts on Human Factors in Computing Systems, pp. 415–420.

Jang, K. L., McCrae, R. R., Angleitner, A., Riemann, R., & Livesley, W. J. (1998). Heritability of facet-level traits in a cross-cultural twin sample: support for a hierarchical model of personality. Journal of personality and social psychology, 74(6), 1556.

Janitzek, D.B. (2012). Gamification: The what and why not!

John Pavlus (November 4, 2010). "Reasons Why "Gamification" is Played Out".

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). The big five inventory—versions 4a and 54. Berkeley: University of California, Berkeley, Institute of Personality and Social Research.

Johnson, D., & Gardner, J. (2010). Personality, motivation and video games (p. 276). ACM Press. doi:10.1145/1952222.1952281

Johnson, D. W., Maruyama, G., Johnson, R., Nelson, D., & Skon, L. (1981). Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analysis. Psychological bulletin, 89(1), 47.

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. Educational researcher, 38(5), 365-379.

Johnson, D. W., & Johnson, R. T. (2005). New developments in social interdependence theory. Genetic, social, and general psychology monographs, 131(4), 285-358.

Johnson, D. W., Johnson, R. T., & Holubec, E. J. (1988). Cooperation in the classroom. Interaction Book Co.

Jolls, C., Sunstein, C. R., & Thaler, R. (1998). A behavioral approach to law and economics. Stanford law review, 1471-1550.

Jones, S., Lynch, P., Maiden, N., & Lindstaedt, S. (2008, September). Use and influence of creative ideas and requirements for a work-integrated learning system. In International Requirements Engineering, 2008. RE'08. 16th IEEE(pp. 289-294). IEEE.

Judge, T. A., & Ilies, R. (2002). Relationship of personality to performance motivation. Journal of Applied Psychology, 87: 797-807.

Judge, T. and Zapata, C. (2014) The person-situation debate revisited: effect of situation strength and trait activation on the validity of the big five personality traits in predicting job performance". Academy of Management Journal [online] amj–2010. available from <http://amj.aom.org/content/early/2014/07/21/amj.2010.0837.short>

Julius, K., and Salo, J. (2013). Designing Gamification. Marketing.

Jung, J. H., Schneider, C., & Valacich, J. (2010). Enhancing the motivational affordance of information systems: The effects of real-time performance feedback and goal setting in group collaboration environments. Management Science, 56(4), 724–742.

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

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. Educational and Psychology Measurement, 20, 141-151.

Kaiser, H. F. (1974). An index of factorial simplicity. Psychometrika, 39, 32-36.

Kani, J., & Nishigaki, M. (2013). Gamified CAPTCHA. In Human Aspects of Information Security, Privacy, and Trust (pp. 39-48). Springer Berlin Heidelberg.

Kaptein, dr M., (2015). Persuasion Profiling: How the internet knows what makes you tick, 6 edition. ed. Business Contact, Amsterdam.

Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. Journal of personality and social psychology, 65(4), 681.

Kelley, H. H., & Stahelski, A. J. (1970). Social interaction basis of cooperators' and competitors' beliefs about others. Journal of personality and social psychology, 16(1), 66.

Kelley, H. H., & Thibaut, J. W. (1954). Experimental studies of group problem solving and process. Handbook of social psychology, 2, 735-785.

Kelley, H. H., & Thibaut, J. W. (1978). Interpersonal relations: A theory of interdependence (p. 341). New York: Wiley.

Kenny, M. E., Blustein, D. L., Haase, R. F., Jackson, J., & Perry, J. C. (2006). Setting the stage: Career development and the student engagement process. Journal of Counseling Psychology, 53(2), 272.

Kerfoot, B. P., Turchin, A., Breydo, E., Gagnon, D., & Conlin, P. R. (2014). An Online Spaced-Education Game Among Clinicians Improves Their Patients' Time to Blood Pressure Control A Randomized Controlled Trial. Circulation: Cardiovascular Quality and Outcomes, 7(3), 468-474.

Kirkcaldy, B., & Furnham, A. (1991). Extraversion, neuroticism, psychoticism and recreational choice. Personality and individual Differences, 12(7), 737-745.

Klehe, U. C., & Anderson, N. (2007). The moderating influence of personality and culture on social loafing in typical versus maximum performance situations. International Journal of Selection and Assessment, 15(2), 250-262.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychological bulletin, 119(2), 254.

Kohn, A. (1999). The trouble with gold stars, incentive plans, A's, praise and other bribes. Boston, MA: Houghton Mifflin.

Koivisto, J., & Hamari, J. (2014). Demographic differences in perceived benefits from gamification. Computers in Human Behavior, 35, 179-188.

Korn, O., Schmidt, A., and Hörz, T. (2012). Assistive systems in production environments: exploring motion recognition and gamification. In Proceedings of the 5th International Conference on PErvasive Technologies Related to Assistive Environments, p. 9.

Korn, O. (2012, June). Industrial playgrounds: how gamification helps to enrich work for elderly or impaired persons in production. In Proceedings of the 4th ACM SIGCHI symposium on Engineering interactive computing systems (pp. 313-316). ACM.

Kotini, I., & Tzelepi, S. (2015). A gamification-based framework for developing learning activities of computational thinking. In Gamification in Education and Business (pp. 219-252). Springer International Publishing.

Kozlowski, S. W. J., & Bell, W. S. (2003). Work groups and teams in organizations. In W. C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds.), Handbook of psychology: Industrial and organizational psychology, Vol 12, pp. 333-375.

Kramer, J., Noronha, S., & Vergo, J. (2000). A user-centered design approach to personalization. Communications of the ACM, 43(8), 44-48.

Kranz, M., Murmann, L., and Michahelles, F. (2013). Research in the Large: Challenges for Large-Scale Mobile Application research-A Case Study about NFC Adoption using Gamification via an App Store. International Journal of Mobile Human Computer Interaction (IJMHCI) 5, 45–61.

Kreft, I. G. (1996). Are multilevel techniques necessary? An overview, including simulation studies.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. Journal of the American Statistical Association, 47(260), 583-621.

Kucklich, J. (2005). Precarious playbour: Modders and the digital games industry. Fibreculture,(5).

Kumar, B. (2012). Gamification in Education-Learn Computer Programming with Fun. Internation Journal of Computers and Distributed systems 2, 46–53.

Kumar, J. (2013). Gamification at work: designing engaging business software(pp. 528-537). Springer Berlin Heidelberg.

Laamarti, F., Eid, M., & El Saddik, A. (2014). An Overview of Serious Games.International Journal of Computer Games Technology, 2014.

Landers, R. N. (2014). Developing a Theory of Gamified Learning Linking Serious Games and Gamification of Learning. Simulation & Gaming, 45(6), 752-768.

Landers, R. N., & Landers, A. K. (2015). An Empirical Test of the Theory of Gamified Learning The Effect of Leaderboards on Time-on-Task and Academic Performance. Simulation & Gaming, 1046878114563662.

Landers, R. N., & Lounsbury, J. W. (2006). An investigation of Big Five and narrow personality traits in relation to Internet usage. Computers in Human Behavior, 22(2), 283–293. doi:10.1016/j.chb.2004.06.001

Landers, R.N., and Callan, R.C. (2011). Casual social games as serious games: The psychology of gamification in undergraduate education and employee training. In Serious Games and Edutainment Applications, (Springer), pp. 399–423.

Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world?. The journal of economic perspectives, 153-174.

Lew, M., Mesch, D., Johnson, D. W., & Johnson, R. (1986). Positive interdependence, academic and collaborative-skills group contingencies, and isolated students. American Educational Research Journal, 23(3), 476-488.

Lewin, K. (1935). A dynamic theory of personality.

Lewin, K. (1936). A dynamic theory of personality: selected papers. The Journal of Nervous and Mental Disease, 84(5), 612-613.

Lewin, K. (1997). Resolving social conflicts and field theory in social science. American Psychological Association.

Liao, H., & Chuang, A. (2004). A multilevel investigation of factors influencing employee service performance and customer outcomes. Academy of Management Journal, 47(1), 41-58.

Lickel, B., Hamilton, D. L., Wieczorkowska, G., Lewis, A., Sherman, S. J., & Uhles, A. N. (2000). Varieties of groups and the perception of group entitativity. Journal of personality and social psychology, 78(2), 223.

Linehan, C., Kirman, B., & Roche, B. (2015). GAMIFICATION AS BEHAVIORAL PSYCHOLOGY 3. The Gameful World: Approaches, Issues, Applications, 81.

Lillienfeld, S. O., Lynn, S. J., Namy, L. L., & Woolf, N. J. (2009). Psychology: From inquiry to understanding . Boston: Pearson/Allyn and Bacon.

Littlepage, G., Schmidt, G., Whisler, E., & Frost, A. (1995). An input- process-output analysis of influence and performance in problem-solving groups. Journal of Personality and Social Psychology, 69, 877-889.

Liu, D., Li, X., & Santhanam, R. (2013). Digital games and beyond: what happens when players compete. MIS Quarterly, 37(1), 111-124.

Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. (1981). Goal setting and task performance: 1969–1980. Psychological bulletin, 90(1), 125.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. American psychologist, 57(9), 705.

Longua, J., DeHart, T., Tennen, H., & Armeli, S. (2009). Personality moderates the interaction between positive and negative daily events predicting negative affect and stress. Journal of Research in Personality, 43(4), 547-555.

Macionis, J. J., & Gerber, L. M. (2011). Sociology: Seventh Canadian Edition.

Maiden, N., Robertson, S., & Robertson, J. (2006). Creative requirements: invention and its role in requirements engineering. In Proceedings of the 28th international conference on Software engineering (pp. 1073–1074). ACM.

Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. Cognitive science, 5(4), 333-369.

Malone, T. W., & Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. Aptitude, learning, and instruction, 3(1987), 223-253.

Maslow, A. H. (1943). A theory of human motivation. Psychological review, 50(4), 370.

Maslow, A. H., Frager, R., & Cox, R. (1970). Motivation and personality (Vol. 2). J. Fadiman, & C. McReynolds (Eds.). New York: Harper & Row.

Matthews, G., Joyner, L., Gilliland, K., Campbell, S. E., Falconer, S., & Huggins, J. (1999). Validation of a comprehensive stress state questionnaire: Towards a state big three. Personality psychology in Europe, 7, 335-350.

Matthews, G., Campbell, S. E., Falconer, S., Joyner, L. A., Huggins, J., Gilliland, K., ... & Warm, J. S. (2002). Fundamental dimensions of subjective state in performance settings: task engagement, distress, and worry. Emotion, 2(4), 315.

McAdams. D. P. (1992). The five-factor model in personality: A critical appraisal. Journal of Personality, 60, 329-361.

McBreen, H. M., & Jack, M. A. (2001). Evaluating humanoid synthetic agents in e-retail applications. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 31(5), 394-405.

McCallum, S. (2012). Gamification and serious games for personalized health. Stud Health Technol Inform 177, 85–96.

McClelland, D. C. (1987). Human motivation. CUP Archive.

McCrae, R. R., Costa Jr, P. T., Ostendorf, F., Angleitner, A., Hřebíčková, M., Avia, M. D & Smith, P. B. (2000). Nature over nurture: temperament, personality, and life span development. Journal of personality and social psychology, 78(1), 173.

McCrae, R. R., & John, O. P. (1998). An introduction to the five-factor model and its applications. Personality: critical concepts in psychology, 60, 295.

McCrae, R. R., & Costa Jr, P. T. (1997). Personality trait structure as a human universal. American psychologist, 52(5), 509.

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. Journal of personality and social psychology, 52(1), 81.

McGonigal, Jane. (2011) Reality is broken: Why games make us better and how they can change the world.

McNamara, J. M., Stephens, P. A., Dall, S. R., & Houston, A. I. (2009). Evolution of trust and trustworthiness: social awareness favours personality differences. Proceedings of the Royal Society of London B: Biological Sciences, 276(1657), 605-613.

Mejia, J. (2013). Impact of Gamification and Shared Situated Displays on Smartphone Application Engagement.

Mekler, E.D., Brühlmann, F., Opwis, K., and Tuch, A.N. (2013). Disassembling gamification: the effects of points and meaning on user motivation and performance. In CHI'13 Extended Abstracts on Human Factors in Computing Systems, pp. 1137–1142.

Melenhorst, A. S., Rogers, W. A., & Caylor, E. C. (2001, October). The use of communication technologies by older adults: exploring the benefits from the user's perspective. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 45, No. 3, pp. 221-225). SAGE Publications.

Mesch, D., Johnson, D. W., & Johnson, R. (1988). Impact of positive interdependence and academic group contingencies on achievement. The Journal of social psychology, 128(3), 345-352.

Michael, D. R., & Chen, S. L. (2005). Serious games: Games that educate, train, and inform. Muska & Lipman/Premier-Trade.

Miller, A. D., & Murdock, T. B. (2007). Modeling latent true scores to determine the utility of aggregate student perceptions as classroom indicators in HLM: The case of classroom goal structures. Contemporary Educational Psychology, 32(1), 83-104.

Mintsioulis, A., & Kristensen, K. K. (2012). Gamification of Social Network Applications (Doctoral dissertation, Thesis from IT-University Copenhagen).

Mollick, E. R., & Rothbard, N. (2013). Mandatory fun: consent, gamification and the impact of games at work. The Wharton School research paper series.

Mollick, E., & Werbach, K. (2015). Gamification and the Enterprise. The Gameful World: Approaches, Issues, Applications, 439.

Moon, J. W., & Kim, Y. G. (2001). Extending the TAM for a World-Wide-Web context. Information & management, 38(4), 217-230.

Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents?. Cognition and Instruction, 19(2), 177-213.

Mowday, R. T., Steers, R. M., & Porter, L. W. (1979). The measurement of organizational commitment. Journal of vocational behavior, 14(2), 224-247.

Mullins-Sweatt, S. N., Jamerson, J. E., Samuel, D. B., Olson, D. R., & Widiger, T. A. (2006). Psychometric properties of an abbreviated instrument of the five-factor model. Assessment, 13(2), 119-137.

Münz, U., Schumm, P., Wiesebrock, A., & Allgöwer, F. (2007). Motivation and learning progress through educational games. Industrial Electronics, IEEE Transactions on, 54(6), 3141-3144.

Murphy, K. R., & Myors, B. (1999). Testing the hypothesis that treatments have negligible effects: Minimum-effect tests in the general linear model. Journal of Applied Psychology, 84(2), 234.

Murray, H. A. (1938). Explorations in personality. New York: Oxford University Press.

Musson, D. M., Sandal, G., & Helmreich, R. L. (2004). Personality characteristics and trait clusters in final stage astronaut selection. Aviation, space, and environmental medicine, 75(4), 342-349.

MWC (2015). Mobile World Congress conference in Barcelona in 2015.

Myers, I. B. (1962). The Myers-Briggs Type Indicator: Manual.

Nacke, L. E., & Mandryk, R. L. (2010). Designing affective games with physiological input. Fun and Games.

Nacke, L. E., Bateman, C., & Mandryk, R. L. (2011, January). BrainHex: Preliminary Results from a Neurobiological Gamer Typology Survey. In ICEC (pp. 288-293).

Neuman, G. A., & Wright, J. (1999). Team effectiveness: Beyond skills and cognitive ability. Journal of Applied Psychology, 84, 376-389.

Neuman, W. L., & Kreuger, L. (2003). Social work research methods: Qualitative and quantitative approaches. Allyn and Bacon.

Nezlek, J. B. (2011). Multilevel modeling for social and personality psychology. SAGE Publications Ltd.

Nguyen, L., Carroll, J., & Swatman, P. A. (2000, January). Supporting and monitoring the creativity of IS personnel during the requirements engineering process. In System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on (pp. 9-pp). IEEE.

Nicholson, S. (2015). A recipe for meaningful gamification. In Gamification in education and business (pp. 1-20). Springer International Publishing.

Nicholson, S. (2013). Two paths to motivation through game design elements: reward-based gamification and meaningful gamification.

Nicholson, S. (2012). Strategies for meaningful gamification: Concepts behind transformative play and participatory museums. Meaningful Play.

Nickel, P., & Spahn, A. (2012). Trust, Discourse Ethics, and Persuasive Technology. Persuasive Technology, 37.

Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: big five correlates of GPA and SAT scores. Journal of personality and social psychology, 93(1), 116.

Nonnecke, B., Preece, J., & Andrews, D. (2004, January). What lurkers and posters think of each other [online community]. In System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference on (pp. 9-pp). IEEE.

Nunes, J. C., & Drèze, X. (2006). The endowed progress effect: How artificial advancement increases effort. Journal of Consumer Research, 32(4), 504-512.

Nuseibeh, B., & Easterbrook, S. (2000). Requirements Engineering: A Roadmap. In Proceedings of the Conference on The Future of Software Engineering (pp. 35–46). New York, NY, USA: ACM. doi:10.1145/336512.336523

O'Connor, M. C., & Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. Personality and Individual differences, 43(5), 971-990.

O'Donovan, S. (2012). Gamification of the games course. Acesso em, 17, 1-8.

Olson, C. M., Rennie, D., Cook, D., Dickersin, K., Flanagin, A., Hogan, J. W., ... & Pace, B. (2002). Publication bias in editorial decision making. Jama, 287(21), 2825-2828.

Ong, M. (2013). Gamification and its effect on employee engagement and performance in a perceptual diagnosis task.

Oinas-Kukkonen, H., & Harjumaa, M. (2009). Persuasive systems design: Key issues, process model, and system features. Communications of the Association for Information Systems, 24(1), 28.

Osborne, J. W. (2010). Improving your data transformations: Applying the Box-Cox transformation. Practical Assessment, Research & Evaluation, 15(12), 1-9

Osheim, D.E. (2013). This Could Be a Game!": Defining Gamification for the Classroom.

Oxford Dictionaries. Oxford University Press. http://www.oxforddictionaries.com/definition/english/productivity (accessed February 23, 2015).

Pardee, R. L. (1990). Motivation Theories of Maslow, Herzberg, McGregor & McClelland. A Literature Review of Selected Theories Dealing with Job Satisfaction and Motivation.

Pates, J., Karageorghis, C. I., Fryer, R., & Maynard, I. (2003). Effects of asynchronous music on flow states and shooting performance among netball players. Psychology of Sport and Exercise, 4(4), 415-427.

Peeters, M. A. G., Van Tuijl, H. F. J. M., Rutte, C. G., & Reyman, I. M. M. J. (2006). Personality and Team Performance: A Meta-Analysis. European Journal of Personality, 20.

Pesce, M, (2015), Contraludics pp 107 in Walz, S. P., & Deterding, S. (2015). The Gameful World: Approaches, Issues, Applications. Mit Press.

Phillips, D. C., & Burbules, N. C. (2000). Postpositivism and educational research. Rowman & Littlefield.

Phirangee, K., & Hewitt, J. (n.d.). Expressing Emotion in an Online Learning Environment Through Nonverbal Cues. Retrieved from http://www.pepperproject.ca/uploads/1/8/9/8/18984529/csse-_may_14_kp_jh.pdf

Pierce, W. D., Cameron, J., Banko, K. M., & So, S. (2012). Positive effects of rewards and performance standards on intrinsic motivation. The Psychological Record, 53(4), 4.

Pinelle, D., Wong, N., Stach, T., & Gutwin, C. (2009, May). Usability heuristics for networked multiplayer games. In Proceedings of the ACM 2009 international conference on Supporting group work (pp. 169-178). ACM.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. Psychological bulletin, 135(2), 322.

Prause, C.R. (2013). Improving the Internal Quality of Software through Reputation-based Gamification. Universitätsbibliothek.

Prause, C. R. (2013). Improving the internal quality of software through reputation-based gamification (Doctoral dissertation, Aachen, Techn. Hochsch., Diss., 2013).

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. Multivariate behavioral research, 42(1), 185-227.

Qin, Z., Johnson, D. W., & Johnson, R. T. (1995). Cooperative versus competitive efforts and problem solving. Review of educational Research, 65(2), 129-143.

Raudenbush, S. W. (2008). Many small groups. In Handbook of multilevel analysis (pp. 207-236). Springer New York.

Reeve, J. (2011). Understanding motivation and emotion. John Wiley & Sons.

Reichheld, F. F. (2006). The ultimate question: Driving good profits and true growth. Harvard Business Press.

Reid-Seiser, H. L., & Fritzsche, B. A. (2001). The usefulness of the NEO PI-R Positive Presentation Management Scale for detecting response distortion in employment contexts. Personality and Individual Differences, 31(4), 639-650.

Reiners, T., Wood, L., Gregory, S., & Teras, H. (2015). Gamification Design Elements in Business Education Simulations.

Rey, P. J. (2015). Gamification and post-fordist capitalism. The Gameful World: Approaches, Issues, Applications, 277.

Richter, G., Raban, D. R., & Rafaeli, S. (2015). Studying Gamification: The Effect of Rewards and Incentives on Motivation. In Gamification in Education and Business (pp. 21-46). Springer International Publishing.

Riemann, R., Angleitner, A., & Strelau, J. (1997). Genetic and Environmental Influences on Personality: A Study of Twins Reared Together Using the Self-and Peer Report NEO-FFI Scales. Journal of Personality, 65(3), 449-475.

Ritterfeld, U., Cody, M., & Vorderer, P. (Eds.). (2009). Serious games: Mechanisms and effects. Routledge.

Robson, C. (1993). Real world research: A resource for social scientists and practitioners-researchers. Massachusetts: Blackwell Pushers.

Ross, S. R., Rausch, M. K., & Canada, K. E. (2003). Competition and cooperation in the five-factor model: Individual differences in achievement orientation. The Journal of psychology, 137(4), 323-337.

Rothbard, N. P., & Wilk, S. L. (2011). Waking up on the right or wrong side of the bed: Start-of-workday mood, work events, employee affect, and performance. Academy of Management Journal, 54(5), 959-980.

Rotter, J. B. (1954). Social learning and clinical psychology.

Roux, A. V. D. (2003). A glossary for multilevel analysis: part II. Epidemiol. bull, 24(4), 10-13.

Rughinis, R. (2013, June). Gamification for productive interaction: Reading and working with the gamification debate in education. In Information Systems and Technologies (CISTI), 2013 8th Iberian Conference on (pp. 1-5). IEEE.

Rupayana, D. D. (2010). Developing SAENS: development and validation of a Student Academic Engagement Scale (SAENS) (Doctoral dissertation, Kansas State University).

Ryan, R. M. & Deci, E. L.(1995). Human autonomy: The basis for true self-esteem. In M. Kernis (Ed.),Efficacy, agency, and self-esteem (pp. 3149). New York: Plenum.

Ryan, R. M. & Deci, E. L. (1991). A motivational approach to self: Integration in personalit'. In R. Dienstbier (Ed.), Nebraska symposium on motivation: Vol. 38. Perspectives on motivation (pp. 237–288). Lincoln: University of Nebraska Press.

Ryan, R. M. & Deci, E. L. (2011). Self-determination theory. Handbook of theories of social psychology, 1, 416-433.

Ryan, R., & Deci, E. (2000a). Intrinsic and extrinsic motivations: Classic definitions and new directions. Contemporary educational psychology, 25(1), 54–67.

Ryan, R. M., & Deci, E. L. (2000b). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist, 55 (1), 68–78.

Ryckman, R. M., Hammer M., Kaczor; L M., & Gold, J. A. (1996). Construction of a Personal Development Competitive Attitude Scale. Journal of Personality Assessment 66, 374-386.

Ryckman, R. M., Hammer, M., Kaczor, L. M., & Gold, J. A. (1990). Construction of a hypemompetitive Attitude Scale. Journal of Personality Assessment 55, 630-639.

Ryckman, R. M., Libby, C. R., van den Borne, B., Gold, J. A., & Lindner, M. A. (1997). Values of hypercompetitive and personal development competitive individuals. Journal of personality assessment, 69(2), 271-283.

Salen, K., & Zimmerman, E. (2004). Rules of play: Game design fundamentals. MIT press.

Sass, M., & Liao-Troth, M. (2015). Personality and Negotiation Performance: The People Matter. Journal of Collective Negotiations.

Saucier, G., & Goldberg, L. R. (2001). Lexical studies of indigenous personality factors: Premises, products, and prospects. Journal of personality, 69(6), 847-879.

Schacter, D. L., Gilbert, D. T., & Wegner, D. M. (2009). Introducing psychology. Macmillan.

Schaffernicht, M. (2006). Detecting and monitoring change in models. System Dynamics Review, 22(1), 73.

Schellens, T., Van Keer, H., & Valcke, M. (2005). The impact of role assignment on knowledge construction in asynchronous discussion groups: A multilevel analysis. Small Group Research, 36, 704-745.

Scherbaum, C. A., & Ferreter, J. M. (2009). Estimating statistical power and required sample sizes for organizational research using multilevel modeling. Organizational Research Methods, 12(2), 347-367.

Schippers, M. C. (2014). Social loafing tendencies and team performance: The compensating effect of agreeableness and conscientiousness. Academy of Management Learning & Education, 13(1), 62-81.

Schuler, H. (Ed.). (2004). Beurteilung und Förderung beruflicher Leistung. Hogrefe Verlag.

Schweder, T., & Hjort, N. L. (2002). Confidence and Likelihood*. Scandinavian Journal of Statistics, 29(2), 309-332.

Scriven, M. S. (1967). The methodology of evaluation (Perspectives of Curriculum Evaluation, and AERA monograph Series on Curriculum Evaluation, No. 1). Chicago: Rand NcNally.

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. International Journal of Human-Computer Studies, 74, 14-31.

Seif El-Nasr, M., Aghabeigi, B., Milam, D., Erfani, M., Lameman, B., Maygoli, H., & Mah, S. (2010, April). Understanding and evaluating cooperative games. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 253-262). ACM.

Seo, Y., Buchanan-Oliver, M., & Fam, K. S. (2015). Advancing research on computer game consumption: A future research agenda. Journal of Consumer Behaviour, 14(6), 353-356.

Seok, S., & DaCosta, B. (2015). Predicting Video Game Behavior An Investigation of the Relationship Between Personality and Mobile Game Play. Games and Culture, 1555412014565640.

Seneviratne, L., & Izquierdo, E. (2010, March). An interactive framework for image annotation through gaming. In Proceedings of the international conference on Multimedia information retrieval (pp. 517-526). ACM.

Shaer, O., Kol, G., Strait, M., Fan, C., Grevet, C., & Elfenbein, S. (2010). G-nome surfer: A tabletop interface for collaborative exploration of genomic data. InCHI'10 international conference on human factors in computing systems(pp. 1427–1436). ACM Press.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). Biometrika, 591-611.

Sharan, S. (1990). Cooperative learning: Theory and research. Praeger Publishers.

Sharpe, L., & Tarrier, N. (1993). Towards a cognitive-behavioural theory of problem gambling. The British Journal of Psychiatry, 162(3), 407–412.

Shebaya, M. (2011). Leadership development: the role of developmental readiness, personality dispositions, and individual values (Doctoral dissertation, Aston University). [online] available from <http://eprints.aston.ac.uk/16039/>

Shepard, L. A. (2005, October). Formative assessment: Caveat emptor. In ETS Invitational Conference, New York, NY.

Sheskin, D. J. (2011). Handbook of parametric and nonparametric statistical procedures (5th ed.). Boca Raton, FL: Chapman & Hall/CRC Press.

Shields, P. M., & Rangarajan, N. (2013). A playbook for research methods: Integrating conceptual frameworks and project management. New Forums Press.

Shneiderman, B. (2000). Creating creativity: user interfaces for supporting innovation. ACM Transactions on Computer-Human Interaction (TOCHI), 7(1), 114-138.

Siang, A. C., & Rao, R. K. (2003). Theories of learning: a computer game perspective. In Multimedia Software Engineering, 2003. Proceedings. Fifth International Symposium on (pp. 239–245). IEEE.

Sicard, M. (2015). Playing the Good Life. In Deterding and Waltz (eds.) (2015). The Gameful World. The MIT Press.

Sicart, M., (2009). The ethics of computer games. Cambridge: The MIT Press.

Sicart, M. (2015). PLAYING THE GOOD LIFE: GAMIFICATION AND ETHICS 8. The Gameful World: Approaches, Issues, Applications, 225.

Siemsen, E., Roth, A. V., & Balasubramanian, S. (2008). How motivation, opportunity, and ability drive knowledge sharing: The constraining-factor model. Journal of Operations Management, 26(3), 426–445. http://doi.org/10.1016/j.jom.2007.09.001

Singer, L., and Schneider, K. (2012). It was a bit of a race: Gamification of version control. In Games and Software Engineering (GAS), 2012 2nd International Workshop On, pp. 5–8.

Singer, J. D., & Willett, J. B. (2003). Applied longitudinal data analysis: Modeling change and event occurrence. Oxford university press.

Skinner, B. F. (1953). Science and human behavior. Simon and Schuster.

Skinner, B. F., & Holland, J. G. (1961). The analysis of behavior: A program for self-instruction. New York: McGraw-Hill.

Skinner, B. F. (2014). Contingencies of reinforcement: A theoretical analysis (Vol. 3). BF Skinner Foundation.

Smith, J. K. (1983). Quantitative versus qualitative research: An attempt to clarify the issue. Educational researcher, 6-13.

Smith, R. (2011, November). The future of work is play: Global shifts suggest rise in productivity games. In Games Innovation Conference (IGIC), 2011 IEEE International (pp. 40-43). IEEE.

Smith, M. (2014), Conversations prior to running of StarQuest at Coventry University.

Snijders, T. A., & Bosker, R. J. (1993). Standard errors and sample sizes for two-level research. Journal of Educational and Behavioral Statistics, 18(3), 237-259.

Snijders, T. B., & Bosker, R. R.(1999). Multilevel analysis: An introduction to basic and advanced multilevel modeling.

Sonnentag, S. (2012). Psychological Detachment From Work During Leisure Time The Benefits of Mentally Disengaging From Work. Current Directions in Psychological Science, 21(2), 114-118.

Sparling, E. I., & Sen, S. (2011). Rating: how difficult is it? In Proceedings of the fifth ACM conference on Recommender systems (pp. 149–156). ACM.

Speer, R., Havasi, C., & Surana, H. (2010, May). Using verbosity: Common sense data from games with a purpose. In Twenty-Third International FLAIRS Conference.

Spence, J. R., & Keeping, L. M. (2010). The impact of non-performance information on ratings of job performance: A policy-capturing approach. Journal of Organizational Behavior, 31(4), 587-608.

Spinath, F. M. (2001). Genetic and environmental influences on achievement motivation and its covariation with personality and intelligence. Personality and temperament: Genetics, evolution, and structure, 1, 11-25.

Staff, F. C. (n.d.). A Q&A With Salman Khan, Founder Of Khan Academy. Retrieved January 7, 2015, from http://live.fastcompany.com/Event/A_QA_With_Salman_Khan

Stanne, M. B., Johnson, D. W., & Johnson, R. T. (1999). Does competition enhance or inhibit motor performance: a meta-analysis. Psychological bulletin, 125(1), 133.

Statistica (2014). Number of apps available in leading app stores 2014 | Statistic. (n.d.). Retrieved January 14, 2015, from http://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/

Statistica (2015). Apple: most popular app store categories 2015 | Statistic. (n.d.). Retrieved January 14, 2015, from http://www.statista.com/statistics/270291/popular-categories-in-the-app-store/

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. Science, 348(6230), 91-94.

Steffensmeier, J. (2008). Situational constraints and personality as antecedents of organizational citizenship behaviors.

Stewart, G. L. (1996). Reward structure as a moderator of the relationship between extraversion and sales performance. Journal of Applied Psychology, 81(6), 619.

Steinke, J. A. (2011). Team conflict and effectiveness in competitive environments (Doctoral dissertation, Wright State University).

Strijbos, J. W., Martens, R. L., Jochems, W. M., & Broers, N. J. (2004). The effect of functional roles on group efficiency using multilevel modeling and content analysis to investigate computer-supported collaboration in small groups. Small Group Research, 35(2), 195-229.

Stott, A., & Neustaedter, C. (2013). Analysis of gamification in education. Surrey, BC, Canada.

Suls, J., Martin, R., & Wheeler, L. (2002). Social comparison: Why, with whom, and with what effect? Current Directions in Psychological Science, 11(5), 159–163.

Sunstein, C. R., & Thaler, R. (2008). Nudge. The politics of libertarian paternalism. New Haven.

Surugiu, T. (2014). Gamification as a means for employee motivation, personal engagement and behavioral outcomes: a gamification system developers' perspective.

Sweetser, P., & Wyeth, P. (2005). GameFlow: a model for evaluating player enjoyment in games. Computers in Entertainment (CIE), 3(3), 3-3.

Taggar, S., & Parkinson, J. (2007). Personality tests in accounting research. Journal of Human Resource Costing & Accounting, 11(2), 122-151.

Taggar, S. (2002). Individual creativity and group ability to utilize individual creative resources: A multilevel model. Academy of Management Journal, 45(2), 315-330.

Taleb, N. N. (2007). The black swan: The impact of the highly improbable. Random House.

Tang, J. H., Chen, M. C., Yang, C. Y., Chung, T. Y., & Lee, Y. A. (2016). Personality traits, interpersonal relationships, online social support, and Facebook addiction. Telematics and Informatics, 33(1), 102-108.

Tauer, J. M., & Harackiewicz, J. M. (2004). The effects of cooperation and competition on intrinsic motivation and performance. Journal of personality and social psychology, 86(6), 849.

Templeton, G. F. (2011). A two-step approach for transforming continuous variables to normal: implications and recommendations for IS research.Commun Assoc Inf Syst, 28(4).

Teng, C. I. (2009). Online game player personality and real-life need fulfillment. International Journal of Cyber Society and Education, 2(2), 39-50.

Teng, C. I. (2008). Personality differences between online game players and nonplayers in a student sample. CyberPsychology & Behavior, 11(2), 232-234.

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. Personnel psychology, 44(4), 703-742.

Thaler, R. H., Sunstein, C. R., & Balz, J. P. (2014). Choice architecture. The Behavioral Foundations of Public Policy.

Thaler, R. H., Sunstein, C. R. (2008). Nudge. The politics of libertarian paternalism. New Haven.

Thom, J., Millen, D., and DiMicco, J. (2012). Removing gamification from an enterprise sns. In Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work, pp. 1067–1070.

Thorndike, E. L. (1912). The measurement of educational products. The School Review, 20(5), 289-299.

Tjosvold, D., & Johnson, D. (2000). Deutsch's theory of cooperation and competition. In Work teams: Past, present and future (pp. 131-155). Springer Netherlands.

Tjosvold, D. (1990). Power in cooperative and competitive organizational contexts. The Journal of social psychology, 130(2), 249-258.

Tranfield, D. R., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. British journal of management, 14, 207-222.

Tutar, H., Altinoz, M., & Cakiroglu, D. (2011). The effects of employee empowerment on achievement motivation and the contextual performance of employees. African Journal of Business Management, 5, 6318-6329.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. Science, 211(4481), 453-458.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. Journal of business, S251-S278.

Tzou, M.-H. (2012). Blurring the boundary between play and work; Gamification and its takeover.

Unger, K., Schwartz, D., & Foucher, J. (2013, October). Increasing Employee Productivity through Gamification and Blended Learning. In World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education (Vol. 2013, No. 1, pp. 2538-2545).

Van Dijk, D., & Kluger, A. N. (2011). Task type as a moderator of positive/negative feedback effects on motivation and performance: A regulatory focus perspective. Journal of Organizational Behavior, 32, 1084-1105. doi: 10.1002/job.725

Van Mierlo, H., & Kleingeld, A. (2010). Goals, strategies, and group performance: Some limits of goal setting in groups. Small Group Research.

Van Scotter, J. R., & Motowidlo, S. J. (1996). Interpersonal facilitation and job dedication as separate facets of contextual performance. Journal of applied psychology, 81(5), 525.

Vargha, A., & Delaney, H. D. (1998). Kruskal-Wallis test and stochastic homogeneity. Journal of Educational and behavioral Statistics, 23(2), 170-192.

Vassileva, J. (2012). Motivating participation in social computing applications: a user modeling perspective. User Modeling and User-Adapted Interaction, 22(1-2), 177–201.

Vegt, N., Visch, V., de Ridder, H., & Vermeeren, A. (2015). Designing Gamification to Guide Competitive and Cooperative Behavior in Teamwork. In Gamification in Education and Business (pp. 513-533). Springer International Publishing.

Veltri, N., Krasnova, H., Baumann, A., & Kalayamthanam, N. (2014). Gender Differences in Online Gaming: A Literature Review.

Venhuizen, N.J., Basile, V., Evang, K., and Bos, J. (2013). Gamification for word sense labeling. In Proc. 10th International Conference on Computational Semantics (IWCS-2013), pp. 397–403.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS quarterly, 425-478.

Versteeg, M. J. J. M. (2013). Ethics & Gamification design: a moral framework for taking responsibility.

Vignovic, J. A., & Thompson, L. F. (2010). Computer-mediated cross-cultural collaboration: attributing communication errors to the person versus the situation. Journal of Applied Psychology, 95(2), 265.

Von Ahn, L. (2006). Games with a purpose. Computer, 39(6), 92-94.

Von Neumann, J., & Morgenstern, O. (2007). Theory of games and economic behavior. Princeton university press.

Vorderer, P., Hartmann, T., & Klimmt, C. (2003, May). Explaining the enjoyment of playing video games: the role of competition. In Proceedings of the second international conference on Entertainment computing (pp. 1-9). Carnegie Mellon University.

Vrasidas, C. (2000). Constructivism versus objectivism: Implications for interaction, course design, and evaluation in distance education. International journal of educational telecommunications, 6(4), 339-362.

Vredenburg, K., Mao, J. Y., Smith, P. W., & Carey, T. (2002, April). A survey of user-centered design practice. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 471-478). ACM.

Vroom, V. H. (1964). Work and motivation.

Walker, J. H., Sproull, L., & Subramani, R. (1994, April). Using a human face in an interface. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 85-91). ACM.

Watson, J. B. (1925). Behaviorism. Transaction Publishers.

Welbourne, J. L. (1999). The impact of perceived entitivity on inconsistency resolution for groups and individuals. Journal of Experimental Social Psychology, 35(5), 481-508.

Werbach, K., & Hunter, D. (2012). For the win: How game thinking can revolutionize your business. Wharton Digital Press.

Wertheimer, A. & Zwolinski, M. (2013). Exploitation, The Stanford Encyclopedia of Philosophy URL = <http://plato.stanford.edu/archives/spr2013/entries/exploitation/>.

White, R. W. (1959). Motivation reconsidered: the concept of competence. Psychological review, 66(5), 297.

Williams, R. B., & Clippinger, C. A. (2002). Aggression, competition and computer games: computer and human opponents. Computers in human behavior, 18(5), 495-506.

Wilson, J. M., Goodman, P. S., & Cronin, M. A. (2007). Group learning. Academy of Management Review, 32(4), 1041-1059.

Witt, M., Scheiner, C., and Robra-Bissantz, S. (2011). Gamification of online idea competitions: Insights from an explorative case. Informatik Schafft Communities 192.

Wittgenstein, L., & Anscombe, G. E. M. (1958). Philosophical investigations (Vol. 255). Oxford: Blackwell.

Woo, Y., & Reeves, T. C. (2007). Meaningful interaction in web-based learning: A social constructivist interpretation. The Internet and higher education, 10(1), 15-25.

Wu, Y. L., & Wang, W. T. (2014). Understanding the Effects of Mobile Gamification on Learning Performance.

Wyatt, S., Frost, L., & Stock, F. G. L. (1934). Incentives in repetitive work: A practical experiment in a factory. Stat. Office.

Xu, Y. (2011). Literature Review on Web Application Gamification and Analytics (CSDL Technical Report 11-05).

Yee, N., Bailenson, J. N., & Rickertsen, K. (2007, April). A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 1-10). ACM.

Yee, N. (2006). Motivations for play in online games. CyberPsychology & behavior, 9(6), 772-775.

Yeung, A., Read, J., & Schmid, S. (2012, November). Students' learning styles and academic performance in first year chemistry. In Proceedings of The Australian Conference on Science and Mathematics Education (formerly UniServe Science Conference) (Vol. 11).

Zammitto, V. L. (2010). Gamers' personality and their gaming preferences (Doctoral dissertation, Communication, Art & Technology: School of Interactive Arts and Technology).

Zeineddine, I. (2012). Gamification and Crowdsourcing as engagement techniques for human rights organizations.

Zeis, C., Waronska, A., & Fuller, R. (2009). Value-added program assessment using nationally standardized tests: Insights into internal validity issues. Journal of Academy of Business and Economics, 9(1), 114-128.

Zhao, J., & Ordóñez de Pablos, P. (2010). Analysis of cooperative mechanism of industry–academy R&D alliance lab and case study. Human Factors and Ergonomics in Manufacturing & Service Industries, 20(2), 123-134.

Zichermann, G., & Cunningham, C. (2011). Gamification by design: Implementing game mechanics in web and mobile apps. " O'Reilly Media, Inc.".

Zichermann, G., & Linder, J. (2010). Game-based marketing: inspire customer loyalty through rewards, challenges, and contests. John Wiley & Sons.

Zyda, M. (2005). From visual simulation to virtual reality to games. Computer, 38(9), 25-32.

7 APPENDICES

APPENDIX 1: SUMMARY OF CODED EMPIRICAL STUDIES OF GAMIFICATION

	Aim/Objective of Study	Method	Conclusions
Berenguer es et al. (2013)	To measure the effect of using emoticons and sounds to motivate more recycling.	During four weeks collection rates of a standard bin vs a bin which rewarded users with an emoticon was measured.	The bin with emoticons collected 3 times more than the standard bin.
Brauner et al. (2013)	To investigate the effects of gamification in promoting and hindering engagement in exergames. Contrasting young vs elderly as well as gamer types, personality factors and technical expertise.	71 volunteers between age 20 to 86. In a controlled experiment were asked to play an exergame controlling for age, inclination towards games (whether the participants liked playing board games, computer games or movement games), gaming motivation and performance variation.	72 % stated that the Exergame would increase their motivation to exercise. Older players were inclined to replay the game more than younger users. Level of performance motivation was not correlated with the Bartle gaming score. Performance within the game is not affected by performance motivation but by gamer type. i.e those who prefer movement games, performed better.
Brewer et al. (2013)	Investigate the effect of gamification on task completion for children between 5 to 7.	Two groups of 7 children mean age 6 yrs participated in a standard and gamified version of the task. The second study included prizes such as stickers, toy cars etc.	Use of gamification increased the study task completion rate from 73% to 97%. Use of points for completed tasks increased engagement. Use of prizes that children earned throughout the study but only received at the end can motivate children to complete tasks in empirical lab designs.
Burkey et al. (2013)	To measure and improve student's interest and engagement toward a course using gamification .	51 and 46 senior students across two years were divided pre- and post gamification of their activity. Measures were taken using experience points, grades attained and pre- and post-surveys to measure their attitude toward the course.	Result of T-test shows there is no significant difference in the grades achieved whether gamification was used or not. From the qualitative survey; a) attitude toward the class was improved with gamification. b) experience points for activity was valued highly by the students. c)team based cooperation was improved.
Cornelisse n, Ferdinand et al. (2012)	To investigate and report the use of applying gaming aspects as a means to promote self-study and increase motivation to train for executing operations on human space flight missions that have a	The study took about 1 and half hours. During the evaluation, a specialist was seated next to the participant to observe and answer any questions. The experiment was performed by fifteen participants who walked through the demonstrator with	Results found that participants didn't engage with the game-like features as first predicted. Many users found the demonstrator too child-like and some even got put off by having gamification aspects like achievements included. More focused research needs to be done in order to conclude any application for gamification in astronaut training.

	duration that exceeds the typical duration of low earth orbit missions.	the test leader.	
Depura and Garg. (2012)	To evaluate the use of gamification for employee orientation.	248 new hires were invited to participate in survey questions pre and post joining a firm.	According to the researchers 92% participated and 75% completed all rounds. 41% named the game format as an appealing method of engaging with the enterprise social networking site.
Duarte (2012)	To study the effects of visualisation and gamification techniques in motivating participants in a requirements elicitation task.	21 participants reviewed three rounds of design and refinement on a platform created for eliciting user requirements.	The research supports the use of gamification as having a positive impact on stimulating user involvement.
Dubois and Tamburrelli (2013)	Measure effects of gamification to engage and motivate software development students.	32 groups of between 2 or 3 students were tested in two cases. A) could only see their own metrics B) could see the metrics of all other groups within the same task.	Overall very little difference between the gamified (using competition) and non-gamified approach. In case A students try to maximize their own metrics, in B students use the metric of other students as a benchmark for their activity. The researcher concludes that competition may be better than no competition.
Dumitrache et al. (2013)	Compare a gamified crowdsource application for extracting annotation from medical text with Natural Language Parser.	11 participants, 10 players engaging with the gamified version and 7 engaging with the simple version.	The quality of annotations through the gamified example where comparable to those of the NLP parser (i.e. gamification did not improve the outcome). Allowing players to access each others' answers increased agreement between annotators but decrease the number of suggested tags and annotations.
Fitz-Walter and Tjondronegoro (2011)	To examine the use of context (where the user is, what they are doing) to trigger game elements for a field study of a mobile application used for orienting new students in the university.	26 first year university students used the application. Their usage pattern together with questionnaires were collected and analyzed.	Overall the reported use of context was well received by participants when compared to game elements that required no context. 50% of users prefered location based QR codes to keypad as input devices (prefered by 7.7%)
Fitz-Walter and Tjondronegoro (2011)	Investigates the use of game achievements when applied to a mobile application designed to help new students at university.	26 first year university students (male = 17, female = 9), ages ranged from 17 to 45 years old with an average age of 20 years. Participants took part in the study during orientation week at university,	The results concluded show that adding gamification elements such as achievements were enjoyable for the students when using the app, however many of the activities in the app such as checking in (82% check in 1-3 times) and adding friends (68% added 1 or less

		they were provided a link to download the application and set it up on their mobile device	friends) were only used 1-3 times.
Foster (2012)	Assess the impact of gamification of an engineering course on the engagement and learning levels of the students. In particular to examine whether a greater number of tasks or examination of tasks in a deeper and more nuanced way is achieved.	From a cohort of 300 students, two groups of 5 individuals were selected representing course entrants from 2010 and 2011. In each year the number of analysis was counted.	The 2010 cohort (no gamification) managed 3 sets of analysis. The 2011 cohort (with gamification) achieved 8 sets of analysis. Suggesting a significant increase in the use of design engineer thinking through use of gamification. Low performing teams showed the greatest increase in evidencing design engineer's perspective.
Frang (2012)	To increase the knowledge of enterprise gamification and provide a proof of concept on its ability to increase usage and employee engagement.	Survey of 15 employees together with literature review.	The employees of the firm surveyed were motivated by self-actualisation but not by competition or status. Therefore any designed game layer was to focus on the individual development not by comparing progress. The researchers noted that the main activity of the game 'should be' to write a log of the employees activity.
Firth (2013)	Investigate the impact of gaming elements on location-based social networks (such as Foursquare) on people's mobility decisions.	36 interviews of frequent Foursquare users. 16 women and 20 men. Using grounded theory approach (Charmaz 2006; Glaser & Strauss 1967) Coding and theming as the interviews took part.	Badges seemed to be enough reward in themselves to encourage behaviour. The reward of gaining 'mayorship' had a profound effect on where participants chose to go and how frequently they would check in.
Gnauk (2012)	To explore the motivational framework of gamification using points and leaderboards to engage customers in managing their electricity demand using their smart grid.	12 subjects were provided with a prototype interface to set review and set their electricity usage. The survey used a likert-system with an AttrakDiff and System Usability Scale.	The study showed a potential high user acceptance and consumer engagement to participate in such a platform.
Hahn (2012)	This research evaluates Ribbon Hero 2 (RH2), a gamified help system for Microsoft Office, and investigates whether it supports and encourages users' learning and mastering of a software	The research consisted of two elements; the evaluation of RH2 using the game flow criteria proposed to clearly show that it is a game despite its claims of the creators and a situated co-inquiry with 6 participants to determine their	Based on the game flow theory of RH2 review several issues conceptually and technically were uncovered. Moreover during the co-inquiry similar issues arise which shows that bugs in a learning environment can cause dramatic impact on motivation and learning outcome. In conclusion, RH2 is not ready for "prime

	application.	attitude towards learning and games.	time", but due to the analysis of the co-inquiry people's perceptions of using games for help software are still positive.
Hamari (2013)	To investigate if adding badges with clear goals and the ability to compare with others to a peer-to-peer trading system increased user activity and sociability.	Experiment (2x2 design) conducted over 1.5years, existing users on the platform were evenly assigned to four test groups; Group one had no social comparison or goals, Group two had social comparison but no clear goals, Group 3 had clear goals but no social comparison and Group 4 had social comparison and clear goals.	There was no support for the two hypotheses;A) social comparison increases usage activity and B) clear goals increase usage activity. The group with badges did not produce statistically significant improvements in the quality of the work. Therefore concluding gamification did not have an impact on user behaviour.
Hamari and Koivisto (2013)	Investigate how social factors (social influence, recognition, reciprocal benefit, network exposure) predict attitude toward gamification and intention to continue using the gamification service.	107 participants were surveyed with respect to their attitudes towards gamification.	59.8% reported their intention to use the service was because of gamification. 45.1% would recommend the system because of gamification. Of the reasons to engage - 13.4% reported social influence. 33% recognition and 44.6% perceived reciprocal benefit. Results indicate 'recognition' is not directly affecting attitudes toward gamification, although it may have had the indirect effect through perceived reciprocal benefits. Enabling users to get exposed to attitudes of others and receive feedback can positively influence attitude - perhaps because of gaols (Locke et al. 1990)
Herzig et al.(2012)	Measure the impact of gamification in improving user participation and engagement across 10 hypotheses along telepresence, interactivity, flow, enjoyment, perceived usefulness, ease of use, and behavioral intention.	112 participants evaluated through Technology Acceptance Model a prototype of a gamification of SAP Enterprise Resource Planning software, compared with existing non-gamified version.	Gamification version improved enjoyment by 53%, flow by 30%. Experiment confirmed : Telepresence has a positive effect on flow. The perceived quality of the interface has a positive effect on flow. Perceived enjoyment has a positive effect on perceived ease of use(PEOU). PEOU has a positive effect on Behaviour Intention. Experiment disproved : Flow has a positive effect on perceived usefulness and perceived ease of use. Speed of interaction has a positive effect on flow.
Herzig et al.(2012)	To present a generic platform for enterprise gamification, deriving an architecture for	Based on a literature review, identifying the requirements many gamified systems need; provide immediate feedback,	The gamification of enterprise systems offers a wide range of relevant topics for researchers and practitioners. Moreover, a generic architecture for gamification and its

	gamification systematically from service-oriented and event-driven principles and best practices.	author goals/challenges and rules to the users, the proposed architecture was created. A prototype is built to demonstrate the basics of the architecture and leaves room for future research.	integration into Enterprise Information Systems (EIS) has been proposed, future research will try to investigate the mobile domain as well as generic front end integration scenarios.
Hilton (2013)	To motivate students to think before answering questions in WebIDE; a web-based online learning environment for teaching Java and C concepts. By redesigning the user interface and adding a scoring mechanism.	Conducted through the first fall quarter of 2012 where two groups; Control were a new user interface was applied but no score mechanic and Experimental were a new interface and score mechanic, were tested. Historical survey data from 2011 was used to compare between the new interface and old interface.	There were no significant statistical differences in the output between the new and old interface. Implementation of a score system improved the attitude of students toward the application however it did not yield in higher performance as compared with the control group.
Hsu et al. (2013)	To identify attractive gamification features for collaborative storytelling website.	A hierarchical system structure of gamification design were conducted through group interview with 18 users to identify 35 gamification features. The relative attractiveness of these features were measured through an online survey of 6333 participants.	The top 10 most attractive gamification features could account for more than 50% of attractiveness amongst the 35 features identified. Unpredictable time pressure was an important feature which seems missing from previous relevant studies. Top 10 features were : 1.Clear relationship between act and reward, 2. Types of time pressure unpredictable. 3. Instructions are easy to learn. 4.Users can build their own team. 5. Badges are diverse. 7. Leaderboards are diverse. 8. Points can be easily accumulated. 10.Gifting is possible.
Hulsebosch (2013)	To identify methods of improving user involvement using gamification by targeting specific audiences by their demographic.	Literature review of the effects of demographics on engagement in games where explored. Afterwards a series of case studies were analysed to see whether the characteristics of the target audience matched user participation in a number of cases.	Men score significantly higher on the Bartle game score for activities pertaining to Achievement - Advancement and Competition. Women score significantly higher on Relationship building and maintaining components.
Jacobs et al. (2013)	To asses if a designed game could provide a positive effect on motivation for	Two stroke patients played the game, providing feedback on IMI (Intrinsic Motivation Inventory) and CEQ	Mean score of 5.8 for IMI, implying a positive result of enjoyment, perceived competence, effort, usefulness and relatedness. CEQ returned similar results as

	rehabilitation training of those who had had a stroke.	(Credibility/Expectancy Questionnaire.	regular therapy. Game was noted as not a replacement for exercise.
Janitzek (2012)	To explore the causes for the lack of take up of gamification - the hypothesis being 'gamification has not diffused the way it should have'	A holistic single case study, interviewing gamification 'experts' and gamification vendors.	The key finding is that the diffusion of gamification is due to the high level of uncertainty around the innovation that the early majority do not see the value of investing in the domain. Since gamification does not require any real skill of the end user it can be perceived as shallow. It is difficult for a gamification application to provide real challenges, that would require the end-user to fail. The dominant design is described as 'Foursquare blueprint' - taking in points, badges, levels and incentives.
Julius and Salo (2013)	To study if the MDA framework can be used in gamification of marketing.	5 participants took part in semi-structured interviews.	MDA framework (Werbach & Hunter 2012) is used by all companies interviewed. However transcript analysis revealed that a strict order of elements presented in the model is not paramount in gamifying.
Kani et al. (2013)	To find out if gamification can be applied to the CAPTCHA system to improve security and usability.	Experiment used 10 volunteers and split into three groups; Group 1 was the text recognition-based CAPTCHA system (control), Group 2 swapping CAPTCHA system and Group 3 deletion CAPTCHA system. Users took a survey after they completed each CAPTCHA test.	Results shows that the questions on the survey; “Enjoyed”, “Happy” and “One-more-time”, they averaged 4 or more points for the Gamified CAPTCHA system. This was in comparison with the text recognition-based CAPTCHA (control). The research concludes that the proposed Gamified CAPTCHA system has a higher entertainment value than traditional CAPTCHA systems.
Korn et al. (2012)	To investigate the effects of gamification in design of assistive systems for physical production in improving workers' motivation and the quality of the work and the product.	A design study considering gamification elements were undertaken. Although this was not tested.	A set of conditions to improve flow were identified. These included being involved in an activity with a clear set of goals. Good balance between perceive challenges and perceived skills. Provide clear and immediate feedback. Provide an intrinsically rewarding experience.
Korn (2012)	To introduce an approach for implementing motivating mechanics from game design to production environments by integrating them in a	An evaluation of a past study which included 134 German companies who are willing to use an enhanced assistive system in production. The study showed about 17%	The research introduces an approach for the integration of the gamification element within the larger framework of ASiP (assistive system in production environments). Motion recognition software allows analysis of the work in

	new kind of computer-based assistive system.	companies already employ more than 6% of impaired workers which was a good basis to create the proposed gamified assistive system "WizMo".	real-time, which is then visualised as bricks in the game.
Kranz et al. (2013)	To investigate whether Gamification can be used in an App Store to increase Near Field Communication (NFC) adoption	The game which was developed for the experiment called NFC heroes was deployed on the Google Play app store, this was to reach a large number of players. Feedback was in the form of the number of app installs and app updates.	Less than 50% of users tend to update their application within 7 days of a new update, leaving them open to security issues.
	To explore whether gamification could increase motivation and engagement in learning programming.	207 students were surveyed about their attitude toward gamification to combine gamification with programming course.	A strong preference for use of gamification in imparting programing knowledge was found. 86% of respondents stated they would like to learn using a fun game - rather than not. 67% noted they would rather be evaluated using game levels instead of assessments, mini projects or class tests.
Landers and Callan (2011)	To evaluate the efficacy of a social networking site with gamification (certification) features in motivating students to complete optional learning tasks.	385 students created profiles on the platforms. 113 (29%) participated in the gamified activity (to earn certification for completing optional multiple choice tests). A mean of 4.4 ranks per students. A qualitative survey of 155 students was collected at the end of semester.	Qualitative survey showed the attitude towards the gamified certification programme was perceived as 'enjoyable' and 'rewarding' and not 'fun'. Researchers concluded that offering rewards within a social context that is meaningful to learners can motivate them to complete optional learning tasks that they would otherwise be unlikely to complete.
McCallum (2012)	To review the usage of a reminiscence style game in dementia sufferers.	User observation of 40 participants through a Functional Independence Measure were carried out.	The researches most important result was that 52% of the players did not require help with the game.
Mejia (2013)	Investigate the use of gamification to encourage sustained engagement within a mobile app and increase use of digital signage in public spaces.	Two versions of the application run successively, participants engaged anonymously	Inconclusive support for gamification, it appears to increase engagement although not clear whether this is due to 'prizes' being offered or as result of game elements.
Mekler et al. (2013)	To examine the effects of points and meaningful	172 participants took part in a 2x2 between-subject design	Whilst points did motivate participants to generate more tags, a meaningful frame

	framing on intrinsic motivation and performance in an online image annotation task.	(points vs no points) (framing vs no framing) their performances measured. This was followed by an Intrinsic Motivation Inventory (Deci et al 1994) on a Likert scale.	inspired them to do better at the task and create more quality tags. Overall, the combination of points and meaningful framing yielded the best results. Both points and meaning on their own and combined increased intrinsic motivation.
Mintsiouli s and Kristensen (2012	A proposal to develop a framework for gamification design in social media applications. Using (Deterding et al 2013) definition of gamification and MDA (Hunicke et al, 2004) approach as a base-line.	The proposed gamification framework will be applied to two case studies; Gruvi, which supplied movie fans with the latests trailers and community fan pages and Scandia Housing, one of Denmarks biggest negotiator of exclusive rentals. Both case studies are identified as social network servicers neither of which are within a game context.	The experiment suggests the first steps on bridging the game between marketing and game design, by suggesting a uniquely modified version of the MDA framework for specific use in gamification. This new framework expands on the MDA concept of aesthetics to include sixteen basic desires, so that they may focus on developing aesthetics experience based on the intrinsic motivations of the gamified artefacts for the intended audience.
Mollick and Rothbard (2013)	To evaluate the relationship between consent to use of gamification within the workplace and the subsequent performance within the gamified intervention. To evaluate whether greater game-play outside of work increases employees consent to gamification at work.	Conducted over eighteen days at a startup enterprise, BigDeal. There were three control conditions; one game condition, and two control conditions. The baseline control condition involved no changes, alternative control condition involved participants being exposed to a leaderboard and the game condition participants were exposed to leaderboard and daily emails to give them performance feedback in the game.	Results show the outcome of the game seemed to have little impact of consent. The winners of the game only marginally were more likely to indicate consent to the game than the losers.
O'donovan (2012)	To determine if gamification (and which elements) could improve the motivation of students to do more coursework.	90 anonymous university and college students participated in two surveys.	The elements found to have the most motivating effect in order were : Progress bars, Badges, Storyline and Leaderboard. Those with least effect were forums and stars.
Ong (2013)	To examine the effect of gamification in a perceptual diagnosis task.	Forty participants completed a 22-minute visual search task. Participants were randomly assigned to four conditions resulting from the factorial combination of narrative mechanics (narrative and	Shows gamification has potential for future applications for the work environment, however perhaps the most interesting finding is that the initial oxygen response and self-reported pre-task motivation have differential predictiveness; the former predicted learning gains, whereas the latter

		control condition) and the points mechanic (Points and no-points condition). Attention effort, motivation, and work engagement were measured through performance metrics, functional near-infrared spectroscopy (fNIRS), and self-report questionnaires.	overall performance.
Osheim (2013)	To define gamification in the context of a college classroom, using Foucault's concept of heterotopia.	Study employs the method of heterotopian rhetorical criticism and the methodology of auto-ethnography to analyze World of Warcraft and re-imagine experiences in the game through critical communication pedagogy	The investigation found that there are three fundamentals to gamification in the classroom; Gamification must consist of high-choice, low-risk engagements in a clearly structured environment.
Prause (2013)	Study presents CollabReview, which address the developers' motivation to invest in internal quality without strict regulation. The goal is to create a realistic image of the benefits and disadvantages of the concept.	Several studies were conducted for this research. The first studies are more concerned with technical aspects, e.g. validity of karma score. Later ones deployed CollabReview in natural environments to study social effects. The investigation consisted of reviews with developers, analysing karma scores and recording developer results.	While the gamification itself and its rewards are not the the central focus of this research, several field experiments had to rely on such games to offer an exchange value for karam points to developers. The success of CollabReview fundamentally depends on the perceived exchange value, but if the exchange value is chosen right, then considerable changes in developer behaviour towards more caring for internal quality can be achieved.
Prause (2012)	To investigate the effects of a reputation system with gamified elements in improving source code quality.	Conducted over four weeks, due not being able to accommodate a controlled experiment it was split into two phases; control and experimental phases. Intervention began on week 3 at the experimental phase.	The implementation of gamification had little influence on the code quality. In fact it appears that inclusion of gamification resulted in a small decrease in the quality of the code produced.
Singer and Schneider (2012)	To examine the effects of inclusion of newsfeed and leaderboard in encourage computer science students to make more frequent commits to a version control.	37 students were entered into the platform which measured their activity on the number of commits to a codebase and used it to rank them within their own team and to compare their team to other teams.	A balance of positive and negative comments were found.

		Participants also received notification as well as given provided with feedback upon reaching specific milestones (for instance 1000 commits). A qualitative study of the experience using a Lightweight Documentation of Experience was also collected.	
Stott and Neustaedter (XX)	To investigate how gamification can be used within the education domain by looking at three applications that utilise gamification in the post-secondary setting. Specially looking at five concepts which exist in games; Freedom to Fail, Rapid Feedback, Progression and Storytelling	Evaluation of three case studies; Intro to Information Studies, Just Press Play and Speculative Design of their effectiveness was conducted by identifying key aspects by reviewing course web pages, journal submission, video interviews, online articles and student work.	Gamification has been seen to be useful in a "best practice" sense. There is no one-size fits all model, it's about identifying intrinsic rewards relative to culture of the local community and building game-like interactions on top of those.
Thom et al. (2012)	Examine patterns of user activity in an enterprise social network service after the removal of a points-based gamification system.	3486 participants of the site who had contributed at least one item of content during the four week analysis period were studied. Measuring the number of photos, lists and comments generated, whilst gamification in place and following the removal of the gamification.	Removal of gamification (point based incentive) system significantly reduced the overall contribution of participants. Suggesting that inclusion of points (extrinsic rewards) influences a segment of the user population to participate more intensely.
Tzou (2012)	To evaluate the effects of gamification on exam results of computer science students. (Case two)	60 students took part in a gamified platform to teach concepts of computer science.	In Case Two : The year in which the 60 students were achieved a pass rate of 91% - the highest historic was a pass rate of 85%. From this observation it may be inferred that gamification improved the pass rate.
Venhuizen et al. (2013)	To present how a “Game with a Purpose” (GWAP) called Wardrobe can obtain gold standard data for word sense disambiguation.	A collection of games which make up “Wardrobe” are presented, each targeting a specific level of linguistic annotation. A gold standard annotation is created by experts for a test set of 115 questions with exactly six answers each, which is what is	The researchers gold-standard tagged a portion of their data that was also used in the GWAP. A comparison yielded promising results, ranging from a precision of 0.88 and recall of 0.83 for relative majority agreement, to a precision of 0.98 and recall of 0.35 for questions that were answered unanimously.

		used to evaluate the answers given by the players.	
Witt et al. (2011)	Explorative analysis of motives and perceptions of game mechanics to foster participation, flow, enjoyment and task involvement in an 'online idea competition'.	Qualitative analysis of responses from 30 participants (28 male and 2 female) completed online survey from around the world.	Usage of 'existing knowledge' and 'curiosity' were identified as the most important motive for participation. The allocation of points or ranking were not perceived as an influential factor to alter activity and only had a minor effect on the motivation to introduce further ideas or level of happiness.
Xu (2011)	To explore the various gamification design thoughts and approaches and examine the most commonly employed game mechanics with respect to their usage and effectiveness.	A comparative review of academic and industry implementations on the effectiveness of applying game mechanics to non-game context.	To move beyond the simplistic gamification (inclusion of points and badges) a greater care needs to be taken to create 'gameful' experiences relying on the motivational affordances of the individual. Taking into consideration; meaning (purpose), mastery and autonomy. To focus on effective player journey with intrinsic rewards preferred over extrinsic.
Zeineddine (2012)	To explore motivations for online engagement using gamification for human rights campaigning.	Two case studies of campaigns that meet the criteria of gamification and crowdsourcing within online human rights campaigns.	Positive impact for applying gamification to raise funds, awareness levels, contributions, participation and interactions.

APPENDIX 2: USER SCENARIOS AND PILOTS

The following are excerpts and outcomes from the workshops and interactions with end users during the requirement elicitation process. The overarching requirements which the experiments' platform aims to facilitate together with storyboards that have been defined as journeys lead to specific functionality requirement.

Users expressed the need to facilitate supporting in collecting and sharing documentation. Users need to continuously filter information and to provide relevant information to inspire them. There exist different types of filters and different types of dimensions of information that can be extracted from information resources like websites and software apps. The filtering process is triggered by the input of the user. The system could offer a kind of a collaborative place to be able to collect relevant material which would foster an open attitude to reconsider existing ideas in a playful way.

User Journey and Functionality

Below, the character 'Alex' is a hypothetical persona who could be viewed as a typical system user. In each instance a paragraph describes his activity followed by statements of functionalities required. Alex a student within a study group, he carries out desk research. He uses the system to help find and share information with his group. He selects interesting and relevant information and decides to share this with his group.

- Functionality to receive alerts when new information is available.
- Functionality to record findings and share with the rest of the group.

Users can collaborative with members of their group.

- Functionality for the system to suggest goals.
- Functionality for users to collaborate asynchronously.

Alex receives points for sharing information and contributing creative ideas. He is currently the top point scorer, which secretly gives him a sense of pride. He also enjoys reading the comments posted by others about his ideas. This feedback encourages him to continue submitting comments and suggestions and gives him a sense of ownership for his work.

Alex learnt a new technique for generating new ideas, he shares this technique. When Alex adds this information, he receives more points, which makes him happy. He is still on the top of the leader board and is gaining a reputation for being an expert in his area.

- Functionality for recognition for contributing findings and ideas.
- Functionality to comment on/rate insights and recommendations added by other group members.
- Functionality to receive rewards/points for commenting on/rating insights added by other group members.
- Functionality to receive rewards by completing goals
- Functionality for team members to view a list of rankings, which displays the names of those who have contributed the most to each area.

Results from the formative evaluation.

The evaluation took part between May and August 2014. Twenty two students took part in a number of one day formative evaluation processes. Rather than treating the evaluation of the platforms as a pure usability test, participants were encouraged to use the platform as a real collaboration tool relating to one of their studies.

All participants filled in research consent forms and completed a 44 item personality questionnaire based on the Big Five. The StarQuest platform was introduced through a short presentation, all three versions of cooperative, competitive and control were tested.

StarQuest is intended to be used by teams who are not necessarily co-located over a period of weeks rather than in short sessions. This meant the piloting phase was highly condensed which lead to some issues arising from the use of the platform, as the goal mechanisms are designed to span days and weeks rather than hours. Despite this, the feedback was very useful in helping to shape the project.

Questionnaires measured the participant's view across acceptability, overall objectives, net promoter scores, thematic focus, cognitive absorptions along curiosity, focused immersion and heightened enjoyment, as well as temporal dissociation, social structure and impacts on mental model building. Additionally a qualitative questionnaire enquired about overall impressions, and users' insight into additional features was also

administered, whose results where then incorporated into the full version of StarQuest used for the experiment.

Questions and domains explored in initial feedback on the design for further refinement.

Qualitative questions for *StarQuest* :

What are your overall impressions of *StarQuest* ?

- Playful and fun interface that's easy to interact with the 'Project Health' concept is intriguing
- Interface is friendly, some functions are nicely developed
- Good tool for keeping oversight. in team evaluations communication is a very frequent problem. this feels like a solution
- Intuitive. Nice to see other peoples suggestions. The taks in the sideline are a bit bulky which makes it look cluttered. I'd like to see more content at once.
- A bit complicated at first sight but by trying you learn how it works. There are some usability issues, like ease of uploading image, some icons seem clickable but are not.
- Wow there is content after a brainstorm session!! we actually did document the process
- Would be really helpful in projects. keeping track of your google result is great.

Do you feel *StarQuest* could contribute to your design process? If so how?

- Yes, it is an easy to use log system, but browsing through old logs was not so simple
- Yes, I would use to keep check on the team's dynamic and 'mood' about the project
- The main strength would be forcing the user to actually document int a more meticulous, chronological and communal way
- When I work with others it could be very helpful since your process is documented and you can see where things went right/wrong. I'd see some

problemswith actually integrating it since it the documenting puts the matching on hold.

- It would help in making a visual and inspirational log
- Yes, since it provided overview and collects all my information in one place
- Yes. in collaboration.
- It helps me to make decisions explicit since I have to decide if they are crucial for the process.

Do you feel *StarQuest* would lead to more sharing of progress? (Everyone answered Yes)

- Yes although you have to be sure all team members are using it, otherwise I'd switch back to email
- Yes if use could also rate our projects (for example quality of the architecture
- I believe so as the inspiration that others send, makes me want to contribute since their enthusiasm ...

How did you feel about the project health or ranking? (note the ranking was not fully implemented)

- It would be nice to give 'Kudos' or 'Karma' to a team member
- For now health is only linked to your digital activity. offline activities do not influence health/rankings

What would you add, remove or change in *StarQuest* ?

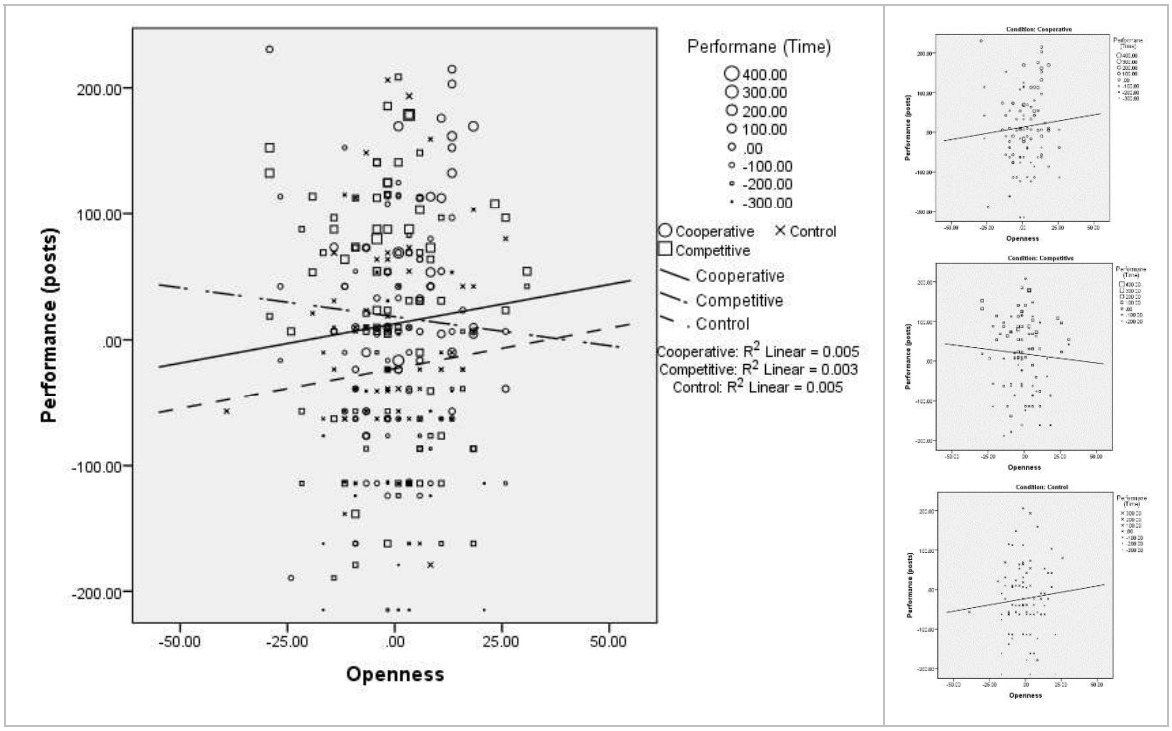
- Filter option in the timeline Assigning tasks to others in the team
- Tagging people who do offline activities. having a 'quality evaluation' to get insight on team's 'mood' and dynamic
- Posting as yourself in the future. complete, one view overview of project as whole. 'Data visualisation' using different variables (tags, type of activity etc). A lot of data being collected, little data being communicated.

- I'd remove 'to do'- as large icons- maybe change it into reminders/may be personalize the visualisation of the team members encourage interaction. Make a quick share button- to minimize time for documenting.
- The act of 'doing something' does not say anything about your contribution to the projects. It would be interesting to also see the value of the contribution and the quality of the project
- More emphasis on future activities - so you can collaboratively plan
- Explore the tool in time on a project combine with time, planning, messages to the future how do posts related to task/ activities?

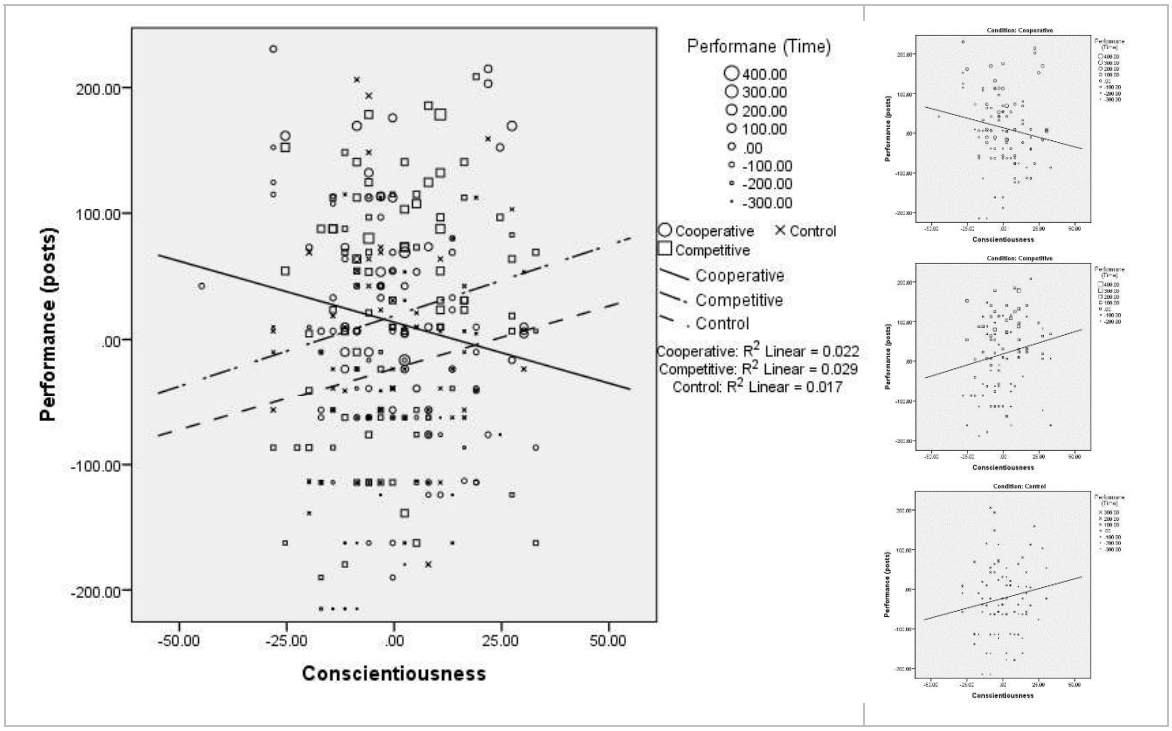
APPENDIX 3 : INDIVIDUAL AND COMPOSITE SCATTERPLOTS AND LINEAR FIT LINES.

The following diagrams present scatterplots of Transformed, Centred Performance Measure *tposts* against each of the Five Personality Traits with tTime represented as size of plot points.

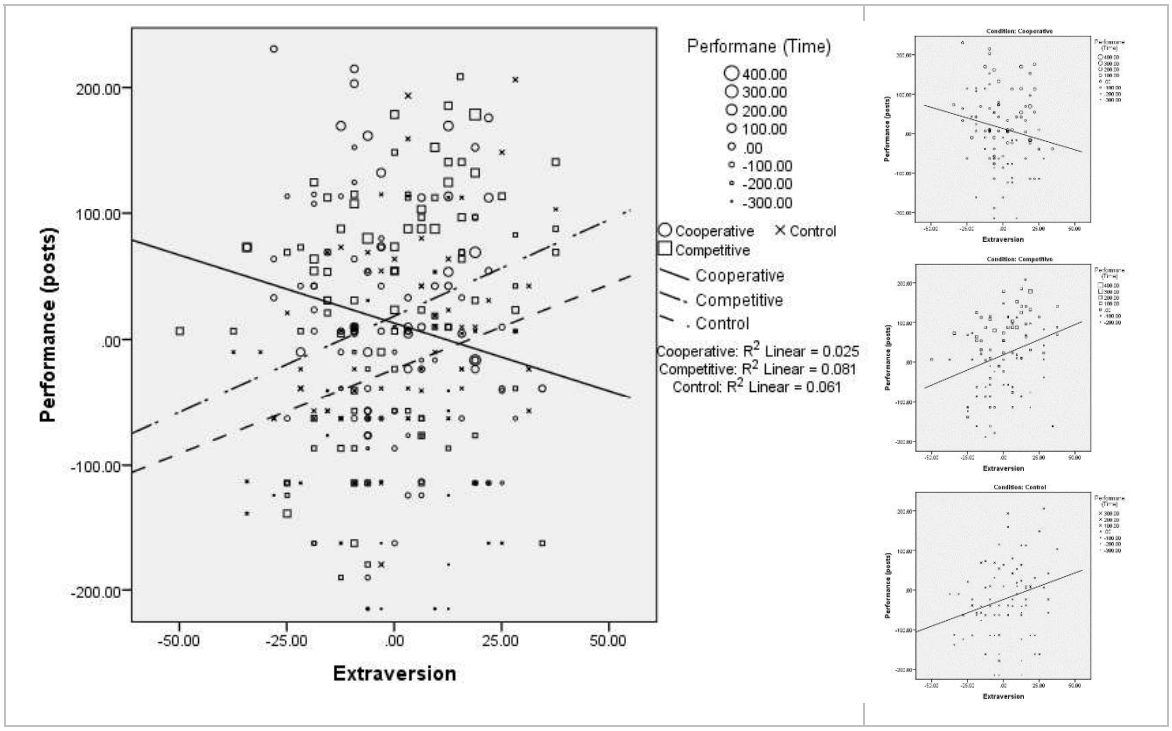
The larger plot on the left is the composite plot of all three conditions, the smaller plots on the right are individual condition plots organised in order from top to bottom of Cooperative, Competitive and Control Conditions.



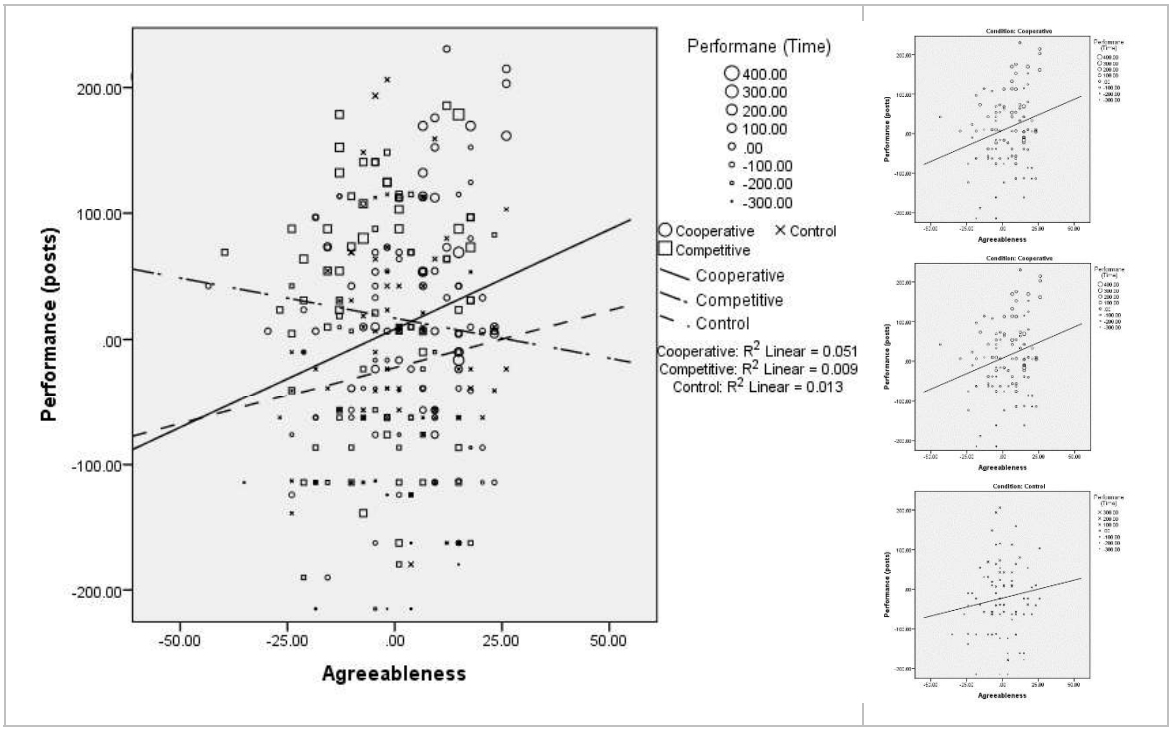
Scatterplots of transformed, centred performance measure *tPosts* against centred Openness with *tTime* represented as size of plot points. Individual condition plots on the right, top to bottom, Cooperative, Competitive and Control condition.



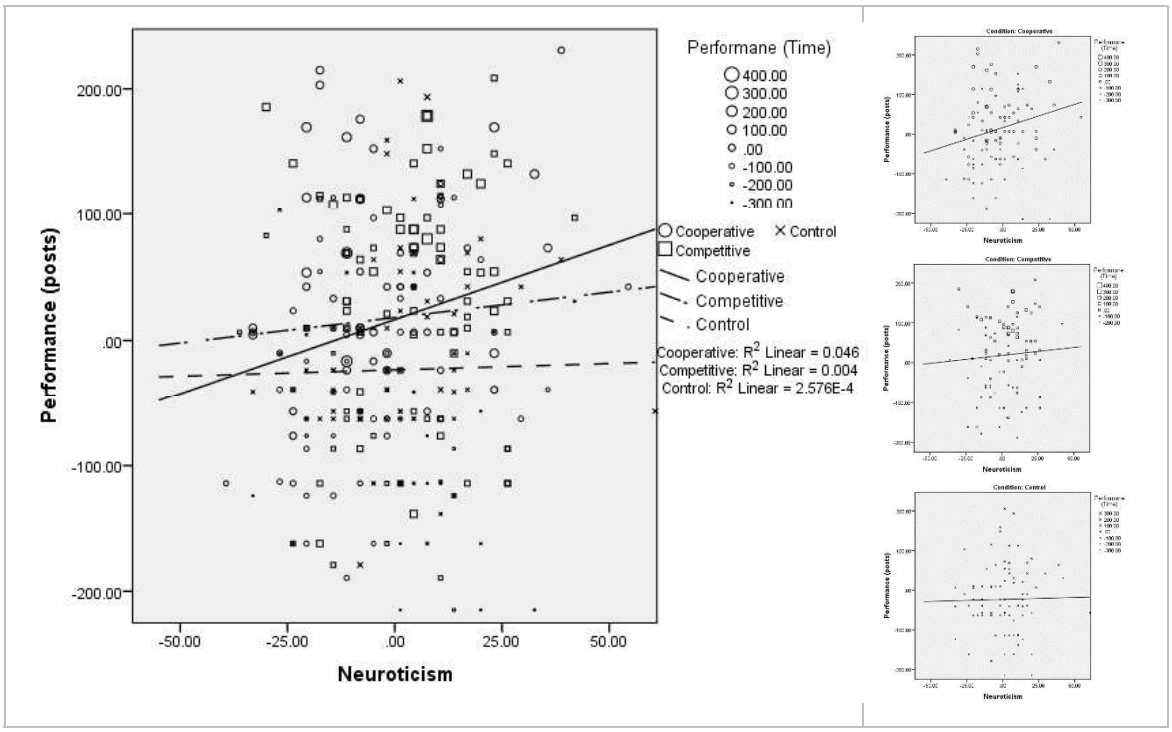
Scatterplots of transformed, centred performance measure *tPosts* against centred Conscientiousness with *tTime* represented as size of plot points. Individual condition plots on the right, top to bottom, Cooperative, Competitive and Control condition.



Scatterplots of transformed, centred performance measure *tPosts* against centred Extraversion with *tTime* represented as size of plot points. Individual condition plots on the right, top to bottom, Cooperative, Competitive and Control condition.



Scatterplots of transformed, centred performance measure *tPosts* against centred Agreeableness with *tTime* represented as size of plot points. Individual condition plots on the right, top to bottom, Cooperative, Competitive and Control condition.



Scatterplots of transformed, centred performance measure *tPosts* against centred Neuroticism with *tTime* represented as size of plot points. Individual condition plots on the right, top to bottom, Cooperative, Competitive and Control condition.

APPENDIX 4: RAW PERFORMANCE SCORES PER COURSE

Raw Performance Scores with Spearman rank order correlations with five factor personality traits

Condition	Course	n	Scores	Mean	Sd	comments	time	Open.	Cons.	Extra.	Agree.	Neuro.
Coop.	C1	64	Posts	11.56	10	.455***	.534***	.076*	-.211**	-.255**	.199*	.109
			Comm	3.3	5.85		.445***	.162*	-.123	-.179*	.259**	.042
			Time	157.3	141.87	.445***		.182*	.069	-.009	.229**	-.064
	C2	22	Posts	7.23	5.282	.484**	.587**	.215	-.403**	-.105	.087	.619***
			Comm	0.14	0.351		.157	.200	-.210	-.231	.326*	.335
			Time	97.9	77.69	.157		.384**	-.132	.121	.324*	.203
	C3	18	Posts	3.17	1.724	.074	.266	.249	-.239	.059	.223	.214
			Comm	0.22	0.548		.318*	.647**	.172	.151	.462**	.118
			Time	95.6	94.37	.318*		.398*	.254	.463**	.675***	-.055
Comp.	C1	66	posts	11.64	7.87	.753***	.675***	-.147	.195*	.309***	-.059	.082
			Comm	3.95	4.72		.657***	-.012	.209**	.156	-.007	.089
			Time	183.6	151.29	.657**		-.074	.111	.005	.006	.181*
	C2	15	posts	8	4.97	.375	.106	.015	.267	.295	-.080	.093
			Comm	0.27	0.594		-.197	-.017	.615***	.504**	-.054	.108
			Time	73.05	33.6	-.197		-.310	-.301	-.619**	-.486**	.567**
	C3	14	posts	3	1.569	-.397*	.792***	.288	-.025	.521**	-.018	-.136
			Comm	0.21	0.426		-.238	.000	.173	-.173	-.110	.111
			Time	34.24	17.98	-.238		.086	.194	.541**	-.326	-.119
Control	C1	59	posts	7.47	7.72	.681***	.633***	.127	.149	.155	.066	.031
			Comm	3.12	4.83		.529***	.160	.138	.104	.124	.136
			Time	135.6	95.7	.529***		.023	.029	-.085	.199*	.009
	C2	17	posts	5.41	4.24	.417**	.228	.098	.365	.433**	.267	-.046
			Comm	0.18	0.529		-.079	.304	.143	.286	.053	.002
			Time	54.04	31.7	-.079		-.083	.122	-.280	.101	.172
	C3	19	posts	3.26	2.38	.039	.590**	.063	-.232	.450**	-.015	-.005
			Comm	0.68	0.946		-.059	.322*	.068	.353*	.047	-.047
			Time	85.68	71.46	-.059		-.195	-.310*	.134	-.220	.034

* p < .1 ** p<.05 *** p<.001

Coop. = Cooperation, Comp. = Competition, Open. = Openness, Cons. = Conscientiousness, Extra. = Extraversion, Agree. = Agreeableness, Neuro. = Neuroticism