Gamification and its effect on employee engagement and performance in a perceptual diagnosis task

---

A dissertation submitted in partial fulfilment of the requirements of the degree of:

Master of Science in Applied Psychology

at the

University of Canterbury

by

Michael Jun Kiat Ong

2013

Research Supervisors:

Dr. William Helton

Paul Russell

Table of Contents

List of Figures

Acknowledgements

First, I would like to thank Deak Helton for all his supervision, direction and support in this research.

Second, I would also like the thank Jonathan Wiltshire for his hard work in creating the experimental program.

Finally, I would like to that my fellow APSY classmates for all their support and company throughout the year.

1. Abstract

Gamification is an emerging phenomenon that has been advocated for its potential to improve organisational outcomes. The present study aimed to examine the effect of gamification in a perceptual diagnosis task. Forty participants completed a 22-minute visual search task. To investigate the role of game mechanics participants were divided into four conditions resulting from the factorial combination of the narrative mechanic (narrative and control condition) and the points mechanic (Points and no-points control condition). Attention effort, motivation, and work engagement were measured through performance metrics, functional near-infrared spectroscopy (fNIRS), and self-report questionnaires. The results revealed points significantly increased task performance while narrative significantly increased intrinsic motivation and prefrontal oxygenation. These findings may provide much needed contributions to the literature surrounding gamification. It was also revealed that fNIRS measures of frontal activation may be a reasonable objective indicator of initial cognitive effort. This presents significant real world applications for objectively measuring motivation.

1. Introduction

2.1 Introduction to Gamification

Every week, across the globe, billions of hours and dollars are spent playing video games. In 2012 the video game industry was valued at approximately $67 billion (Medland, 2013) making it one of the world's largest forms of entertainment. People possess an incredible drive for gaming despite the fact that games are typically played during one's free time, at one's own expense and often involve performing repetitive or menial tasks. 'Gamification' is the idea of harnessing this psychological predisposition to engage in gaming, using the same motivational mechanisms that game designers have become familiar with over the last three decades of making video games, as a potential means to make real world activities more engaging.

An exemplar of this concept is Nike+: Making Fitness Fun (http://nikeplus.com/). Nike+ is a social running game-like application as an answer to the main problem that fitness programmes face – motivation. By employing clever use of game mechanics, Nike+ turns what is essentially a pedometer and stopwatch into something that is fun, social, and engaging. The application tracks the users' activities and allows runners to monitor their progress over time, set personal goals, earn achievements, accumulate points, and gain levels. Exercise data can be easily uploaded to the Nike+ website allowing the user to challenge their friends via online leaderboards and to share their rewards and success stories with each other.

Researchers have defined gamification as "the use of game design elements in non-game contexts" (Deterding, Dixon, Khaled & Nacke, 2011). This definition is intentionally broad so as to encompass its application to a wide variety of areas such as innovation, education and training, employee performance, healthcare, social change, and business and work planning etc. The term 'gamification' only recently came into widespread usage in

2010 (Zichermann & Cunningham, 2011; Google Trends, 2012) but it has since emerged as a recognisable trend. Indeed, the gamification of work is predicted to become more prevalent; (Gartner Enterprise Architecture Summit, 2011). By 2014, Market Researcher, M2 Research, estimates that gamification software, consulting, and marketing revenue will grow to US$938 million from less than US$100 million in 2011.

Many people find the prospect of gamification highly appealing. One study by Saatchi & Saatchi (2011) found that of U.S. residents who were employed, 55% of them were very interested in working for a company that utilises gamification as a method to increase productivity. This number jumps to 83% if you include those who were at least 'somewhat interested'. Video games first appeared in the 1950s and have been mainstream since the late 1970s; the salient question is - why only now is it that gamification is becoming a phenomenon? One of the reasons that has been attributed to gamification's sudden rise in popularity is the changing demographics of the business environment (Burke & Hiltbrand, 2011). For decades, the business workforce has consisted primarily of baby boomers (those born between 1946 and 1964). However, as this generation nears retirement age, Generation X (1965-1978) and Generation Y (1979-2000) will make up the majority of the workforce. These two generations have grown up with videogames in their everyday lives and have become familiar with the platforms that they operate on. The age and age range of people who participate in gaming is increasing; studies show that the mean age of gamers is 35 years (with 12 years of playing video games) and that 26% of gamers are over 50, an increase from 9% in 1999 (Reeves & Read, 2009).

Companies have been using digital games for a number of years to help market their products to consumers and build brand loyalty, but gamification of the workplace, also known as 'enterprise gamification' is starting to gain traction (Silverman, 2011). The proclaimed objectives of gamification include higher levels of engagement, behavioural changes and stimulated innovation (Gartner, 2011); this has obvious potential for practically any area of business application. Indeed, gamification has been successfully demonstrated in a diverse range of industries, for instance, LiveOps Inc., a company which runs virtual call centres, began awarding agents with virtual badges and points for tasks such as keeping calls brief and closing sales. In addition, leaderboards were implemented to allow agents to compare their achievements with others. The gamified system was met with an 80% adoption rate in the very first week that it was introduced. Those who adopted the system were found to have 23% improved performance over non-users, with an average +9% higher rate of customer satisfaction. Moreover, gamification reduced training to 14 hours, down from an average of four weeks (Bourque, 2012).

Another example of game design elements being used to great effect is when the US Navy implemented the Flooding Control Trainer (FTC; Hussain et al., 2009) into its recruitment training programme. The FTC provided training of damage control skills to individuals within the simulated interior of an Arleigh-Burke class destroyer. Utilising a comprehensive narrative arc and a variety of feedback mechanisms, the programme resulted in recruits having a 50% decrease in decision making errors, an up to 80% decrease in communication errors, and a 50% improvement in situational awareness and navigation skills when compared to the control group.

There are numerous other accounts of organisations using gamification that has resulted in greatly improved business outputs (see Herger, 2011). However, the area is primarily driven by practitioners; what is missing is a theoretical model for gamification and the appropriate empirical validation (Herzig, Strahringer & Ameling, 2012).

2.2 Game mechanics

Game mechanics are the constructs of what makes a game a game; they are the tools used in a gamified system that, when utilised correctly, create engaging (but not necessarily fun) experiences. Game mechanics have become the foundation for understanding the psychological drive of games. Some game mechanics have been used for centuries (e.g. badges/titles in the military) while others have been invented with the recent advent of video-games. There are many different types of game mechanics and game elements, SCVNGR, which is a social location-based platform has a playdeck of almost 50 different game elements that their employees are instructed to memorise (Shronfeld, 2010). Leaderboards, achievements, badges, levels, and ranks are all examples of common mechanics. In the present study we will be narrowing the focus to two core game mechanics: points and narrative.

Points: Points are a running numerical value given for any single action or combination of actions. Points are awarded to desired behaviours and are used to give rapid frequent and clear feedback to the user. The cumulative nature of points drives users to continue to remain active.

Narrative: Narrative is the story telling aspect of the game that is central to immersing users into the experience. A game's narrative is built using narrative elements such as theme, backstory, dialog etc. to give context to an event in a game thereby making the activity of playing the game more engaging and less abstract.

Each game mechanics satisfies players' different needs and drive behaviours in various different ways (Xu, 2012). However, theories surrounding the motivation behind games have been mostly founded by game designers and advocates from outside of the mainstream empirical literature. In order to fully understand why games are so engaging it is necessary to understand the fundamental motivational and psychological dynamics that underlie game mechanics (Ryan, Rigby & Przybylsi, 2006).

2.3 The psychology behind gamification

A lot of research has been done in recent years revealing the ability of games to improve performance. Yet the literature regarding the underlying mechanisms behind this process is surprisingly scarce.

Curtiss (2011) highlighted the significant parallel between the basic tenants of game design and the laws of learning. These laws of learning (also referred to as the principles of learning) have been presented as an explanation for why gamification can show such incredible results in performance.

The **principle of readiness** is based on the idea that an unmotivated student is unlikely to learn very much, learning is most effective when the student is mentally willing and physically able. Studies show that motivated students will not only learn more, they will learn at a quicker rate and retain the information for a longer period (Bransford, Brown, Cocking, 2000; Pink, 2009; Williams & Williams, 2010). While many successful games are designed in a way to emotionally entice players into captivating storylines, real-world activities are seldom compelling and it is not always easy for employees to see how their small tasks fit into the larger picture of achieving business success. Gamification uses narrative to build a 'non-fiction' experience into the job so that employees are not just informed of what they need to do but informed about the bigger purpose of why they are doing it. By highlighting and reminding the employee of the impact of their tasks and the importance of their work they are more likely to view their work as meaningful.

The principle of exercise adheres to the idea that our ability to recall information is strengthened by way of repetition, that the mind can rarely recall new concepts from a single exposure. Learning occurs best and is retained longer through meaningful practise and repetition. The law emphasises that learning is only meaningful when practice occurs with feedback. Games excel at employing regular and prompt short and long-term feedback; in contrast, typical feedback loops in the workplace such as annual appraisals are notoriously slow.

The principle of effect is concerned with the emotional reaction from the student. The principle states that when associated with positive emotions such as accomplishment or enjoyment, learning is reinforced. Conversely, learning is diminished when accompanied by negative emotions.

Humans find games inherently fun, a study published by Koepp et al. (1998) found that playing video games triggers the release of dopamine, the same neurotransmitter that signals pleasure when eating food, having sex or taking addictive drugs. Games are excellent at rewarding players correct behaviours, they do this though game mechanics such as points or achievements giving instantaneous feedback. This positive reinforcement motivates the player to learn and succeed through operant conditioning.

The principle of intensity is based on the idea that things that are exciting and engaging are more likely to stimulate learning. Gamification has the potential to make otherwise menial or routine tasks like data entry more exciting and engaging through the use of game mechanics, game design and game thinking. Games are outstanding in producing immersive and engaging experiences and are excellent at inducing the state of flow.

Psychologist Mihály Csíkszentmihályi introduced the concept known as ‘flow’. Flow is a mental state of complete absorption characterised by feelings of full involvement, and success in the process of the activity (Csíkszentmihályi, 1990). One of the critical factors of games is the capability to adjust the balance between ability level and challenge, which is a constituent element of flow. In a gamified system, game mechanics such as points, achievements or objectives can be calibrated so that they are challenging yet achievable within the individual's ability. According to Csíkszentmihályi, flow is excellent at harnessing the emotions that promote performance and learning. Activities that are optimally challenging have been shown to be highly intrinsically motivating (Danner & Lonky, 1981; Gagné & Deci, 2005). Csíkszentmihályi and Rathunde (1993) found that flow is strongly correlated with learning which is unsurprising as the outcomes of flow such as motivation, performance and satisfaction directly link to Edward Thorndike's (1932) first three laws of learning (readiness, exercise and effect).

In a study aimed at trying to understand why games work with grounded psychological theory Ryan and colleagues (2006) proposed that human motivation to engage in video games and the impact of gameplay on motivation and psychological well-being can be accounted for under the framework of Self Determination Theory (SDT). SDT is a widely researched meta-theory of motivation and personality, the theory postulates that the satiation of the basic psychological needs provides the nutriments for intrinsic motivation and internalisation (Gagné & Deci, 2005): The three psychological needs that form the basis of SDT are: autonomy, the universal urge to be have control of one's own life and act with from one’s own volition; competence, the need to develop mastery and take control; and relatedness, the universal want for association, to empathise with and relate to other people.

According to SDT, motivation is not a unitary phenomenon and it is generally distinguished between two different aspects of motivation; intrinsic (internal) motivation and extrinsic (external) motivation (Alexander, Ryan & Deci, 2000). Intrinsic motivation refers a motivation to do a task because they find it inherently interesting or enjoyable; it is a drive that exists within the individual rather than relying on any external pressure. Extrinsic motivation, contrary to intrinsic motivation, is driven by external motivators; satisfaction does not come from the activity itself but rather from the extrinsic consequences of the activity (Deci & Ryan, 1985; Gagné & Deci, 2005). Intrinsic motivation is understood as the core type of motivation underlying videogame participation since they are typically played for the pleasure in the activity being performed rather than for achieving some external reward (Ryan et al., 2006). Indeed, participating in games can be costly in both time and money and may be met with social disapproval. Yet it is possible for games to also be extrinsically motivating; even if someone gains no satisfaction from playing the game itself they may still be motivated to play because of the outcomes, such as praise from their peers or status/rewards that are tied to good performance.

Ryan et al. (2006) found that games are intrinsically motivating because they are structured and designed in such a way that they satisfy the basic needs for autonomy, competence, and relatedness. In the case of the two game mechanics under present examination: performance feedback from the mechanic of points facilitates intrinsic motivation by promoting a sense of competence (Fisher, 1978; Ryan, 1982); while increased interest or personal value from the narrative mechanic can foster intrinsic motivation through greater perceived autonomy.

2.4 Engagement

One of the key reasons that organisations are interested in gamification is its ability to increase engagement. Engaged employees are highly desirable as they are more energetically and effectively connected to their work activities than unengaged workers, and also rate themselves as better able to deal with job demands. Work engagement can be defined as a positive and fulfilling state of mind that is characterised by three key components: vigour, dedication, and absorption (Schaufeli, Salanova, Gonzalez-Romá, & Bakker, 2002). Vigour is characterised by positive levels of energy, mental resilience, and willingness to expend effort on the job, as well as perseverance through difficulty. Dedication is characterised by strong levels of involvement in work tasks accompanied by a feelings of significance, enthusiasm, inspiration, pride, and challenge. Finally, absorption refers to being completely concentrated and deeply engrossed at work, so that time swiftly passes when completing tasks, finding difficulty emotionally detaching from work (Schaufeli et al., 2002).

Engagement is considered to be the antipode of burnout with the core engagement dimensions of vigour and dedications considered the direct opposites of the core burnout dimensions of exhaustion and cynicism respectively (Maslach, Schaufeli, & Leiter, 2001). The relationship between engagement and performance is substantial. A meta-analysis by Harter, Schmidt, Killham and Agrawal (2009) found that employee engagement is related to each of the nine performance outcomes studied: customer ratings, profitability, productivity, turnover (in both high-turnover and low-turnover organisations), safety incidents, shrinkage, absenteeism, patient safety incidents and quality. The results were consistent across different organisations, industries and countries which indicate high generalisability.

The process behind how gamification can increase work engagement can be explained through the job demands-resource (JD-R) model (Demerouti et al., 2001a, b; Bakker & Demerouti, 2007). Motivational resources at the task level, including autonomy, feedback, and task significance have the potential to satisfy the basic human needs as described by SDT; the same psychological needs that Ryan and colleagues (2006) found that games satisfy, which in turn, increases intrinsic motivation (Deci & Ryan, 1985). Job resources may also be instrumental in achieving work goals and thereby can increase motivation extrinsically. According to the JD-R model this increased motivation leads to organisational outcomes of high work engagement, low cynicism, and excellent performance (Bakker & Demerouti, 2007).

2.5 Motivation

In the field of organisational psychology work motivation is the most prominent area of interest for both psychologists and business scholars alike (Steers, Mowday, & Shaprio, 2004; Wright, 2001). Work motivation can be defined as “the force that drives people to behave in a way that energises, directs, and sustains their work behaviour” (Steers et al., 2004). There is good reason that this area to receive such great attention; research has indicated that work motivation leads to increased performance, psychological well-being, organisational trust and commitment, and job satisfaction (Gagné & Deci, 2005).

Matthews, Campbell & Falconer (2001) identified that the two aspects of motivation, Interest (intrinsic) Motivation and Success (extrinsic) Motivation, were both components to a higher-order factor of Task Engagement. Task Engagement related to higher ratings of task demand, effort and performance (Matthews et al., 1999). Designing the work environment so that effective performance leads to both intrinsic and extrinsic rewards has been advocated as a way to produce total job satisfaction (Porter & Lawler, 1968).

Motivation is a subjective state that is most commonly and easily measured through subjective self-reports. While self-reports are the traditional method of assessing motivation in research Matthews et al. (2002) warn that it is potentially dangerous to solely rely on them as they are open to bias by response styles such as acquiescence and social desirability. Moreover, in regards to measuring subjective states though self-reports Matthews et al. (2002) noted that self-reports may only be scratching the surface of measuring emotional states:

> Typically, theories claim that conscious feeling states like motivation represent only a partial expression of some larger psychological system and that a questionnaire scale does not measure the emotion directly, but only indirectly through its expression into consciousness. States are outcome variables, to be distinguished from underlying cognitive and neural processes that may determine state but are not fully accessible to consciousness (pp. 136).

2.6 Functional Near-Infrared Spectroscopy.

In addition to subjective measures, emotional states can be also be accessed through physiological measures. Cognitive demands have been known to interact with psychophysiological and neurophysiological responses in relatively predictive ways

(Fairclough, Venables & Tattersall, 2005). Measuring emotional states though objective measures avert the potential limitations when relying exclusively on subjective measures like self-reports. For instance, since the present experiment involves rescuing victims, even though responses were anonymous and confidential, there is still a very real possibility that participants may over-report their levels of motivation to do the task due to social desirability bias. Furthermore, individuals completing self-report scales are required to not only recall but to also weigh, inference, predict, interpret, and evaluate their state within the experiment. Therefore, any data obtained through such a measure is already somewhat removed from discrete stimuli and responses (Podsakoff & Organ, 1986). Direct neuroimaging techniques such as electroencephalography (EEG) and event-related brain potentials (ERPs) are able to continuously and unobtrusively measure brain functions with the task and do not require individuals to make abstractions of their subjective state. Functional near-infrared spectroscopy (fNIRS) is a sensitive and reliable method of measuring brain activity which has been shown to perform similarly with other functional imaging methods such as fMRI (Huppert, Hoge, Diamond, Franceschini & Boas, 2006). The most commonly used form of fNIRS relies on the optical absorption properties of blood to measure the cerebral haemodynamic response. Light is introduced into the scalp to measure regional changes in blood oxygenation as oxy-hemoglobin (HbO) converts to deoxy-hemoglobin (HbR) during neural activity. fNIRS has been identified as an ideal tool for mental state evaluation due to its high portability, and non-invasiveness (Ishii et al, 2008; Ayaz, et al., 2012) and shows good test-retest reliability for task-specific brain activation (Plichta et al., 2007).

2.7 The Current Experiment

The purpose of this study is to examine whether the application of the game mechanics of points and narrative increases performance, motivation, and work engagement on a perceptual diagnosis task. Advances in computer technologies, as well as the increasing prevalence of information-processing automation have led to an increased importance on perceptual diagnosis (e.g. monitoring) in human-computer interaction (Parasuram & Wickens, 2008). Human analysis and judgment is still needed in systems where automation is not perfectly reliable (Parasuram & Wickens, 2008); this is especially true for safety-critical systems such as air traffic control and intensive care units where the outcomes involve lethality or human safety. For jobs where perceptual diagnosis is a critical skill, it is important to understand how perceptual learning can be improved.

For the present study a visual search task was created to assess participant's diagnostic performance; participants took on the role of an emergency-response worker and were tasked to monitor and detect whether or not images contained buried avalanche victims. In an avalanche, emergency search and rescue workers are usually the only hope for a buried victim as movement is impossible once the snow has settled. The greatest danger in an avalanche is not trauma from the avalanche itself but from suffocation after the snow refreezes. It is therefore crucial for rescuers to work rapidly and efficiently to locate and excavate victims before they succumb to asphyxiation; research in Switzerland based on 442 buried skiers show that survival rates rapidly drop from 92% in the first 15 minutes to 30% after 35 minutes (Falk, Brugger & Adler-Kastner, 1994). Even slight improvement in performance in such emergency rescue systems can mean the difference between life and death.

In the case of vision, repeated exposure to visual stimuli improves the information processing of those stimuli (Mukai et al. 2011; Watanabe et al. 2002). Mere exposure, however, does not guarantee improvement in performance (Mukai et al. 2007). Indeed, variations in attention directed to the stimuli may explain individual differences in learning rates. While some studies have demonstrated perceptual learning occurring to unattended stimuli (Watanabe et al. 2002; 2001), other studies show greater performance improvements in those who appear more attentive and exhibiting signs of greater cognitive effort in regards to the stimuli (Mukai et al. 2007; 2011).

In a previous study, for example, Mukai and colleagues (2007) found that participants who demonstrate perceptual learning initially had higher levels of activation in dorsal fronto-parietal areas in a functional magnetic resonance imaging study. Initial learning places demands on endogenous attention and the executive system, and this is indicated by more frontal cerebral activity.

The literature on the importance of performance feedback for perceptual learning is mixed. Previous research has found that, while repetition and practice is necessary for perceptual learning performance, there is no evidence that perceptual learning requires performance feedback (Gibson & Gibson, 1955). However, other studies have found that feedback improved learning rates and overall performance (Herzog & Fahle, 1998).

Jobs that involve perceptual diagnosis tasks lend themselves particularly well to gamification; as video games have been shown to enhance a wide variety of perceptual and cognitive abilities, such as improved spatial skills, visual attention, and reaction times (Green & Bavelier, 2006; Greenfield, DeWinstanley, Kilpatrick & Kaye, 1994; Castel, Pratt, Drummond, 2005).

In the present study, we investigated the effect of gamification in a perceptual diagnosis task by having participants detect images of snow for buried bodies (see figure 1). To investigate the role of game mechanics participants were divided into four conditions resulting from the factorial combination of game mechanic assignment of narrative (narrative and control condition) and performance points (points and no-points control condition). In order to gage attention effort we employed both functional near-infrared spectroscopy (fNIRS) to track cerebral activity in the frontal cortex and self-reports of pre-task motivation. In addition, post-task self-reports of motivation and work engagement were administered.

We expected performance metrics which indicate perceptual learning, to improve with task exposure. We expected this to be influenced by the game mechanics of points and narrative. In both cases, these may improve performance through improved learning processes and by increasing participants' motivation which may in turn lead to an increase in participant attention effort directed to the stimuli. In addition, we expected an increase in cerebral activity in the frontal cortex early in task exposure and that this activity would decline with increased task exposure indicative of a shift from endogenous effort towards automaticity.

2.8 Predictions

Hypothesis 1 – performance metrics will improve over exposure to the task.

Information processing of stimuli has been found to improve with repeated exposure (Mukai et al., 2011; Watanabe et al, 2002). Consequently, performance metrics, which indicate perceptual learning, are expected to improve with task exposure.

Hypothesis 2 – initial cerebral oxygenation response (in Block 1 of the task) and pre-task motivation will be predicative of who performs better and learns faster.

Improvement in performance has been demonstrated by those who show more attentiveness and exhibit signs of greater cognitive effort in regards to stimuli (Mukai et al., 2007; 2011). It is hypothesised that participants with initially greater frontal cerebral activation and with elevated pre-task self-reported motivation will have more rapid improvement in performance.

Hypothesis 3 – Participants will have higher levels of performance in both game mechanic conditions compared to the control conditions.

Performance is expected to be influenced by the points and narrative game mechanics. In both cases, these may improve performance by increasing participants' motivation and this may in turn lead to an increase in participant attention effort directed to stimuli (Matthews & Davies, 1998; Matthews et al,. 1990; Ryan, Rigby & Przybylski, 2006).

Hypothesis 4 – Participants' levels of work engagement will increase in both game mechanic conditions compared to the control conditions.

Game play has been shown to satisfy the psychological needs of autonomy, competence, and relatedness (Ryan, Rigby & Przybylski, 2006). This is expected to instigate a motivational process leading to greater levels of work engagement. (Bakker & Demerouti, 2006).

3. Method

3.1 Participants

The study was conducted with 40 participants (12 male, 28 female). The mean age of participants was 24.43 years with ages ranging from 18 to 52 years (SD = 6.63 years). Participants were recruited from the University of Canterbury campus via flyers on noticeboards and the University's Psychology research forum on Facebook. All participants were right handed and had normal or corrected-to-normal vision. Participants were reimbursed for their time with a $15 NZD Westfield mall shopping voucher. The study was reviewed and approved by the University of Canterbury's Human Ethics Committee prior to commencing this research.

3.2 Materials

3.2.1Task

The main task of the experiment involved a feature present search task which served to measure performance. During the task participants observed the repetitive presentation of images of snowfields from an aerial viewpoint. The images were presented for 2000ms in which the critical stimulus was either present (rarely occurring) or absent (frequently occurring) as seen in figure 1.

In the task, the critical stimulus was the presence of a faint silhouette of one of three human shaped bodies buried in the snow as seen in figure 2. The body stimuli were approximately 45mm to 65mm in length on the screen. The overlaying snow photograph stimuli was set at 45% transparency to allow the grey body stimuli to look as though they were underneath the snow. The bodies were presented in a random location and orientation on the snowfield image. The order of image presentation was random except for the requirement that the critical stimuli appeared with a probability of $p = 0.\overline{09}$ for each of the four

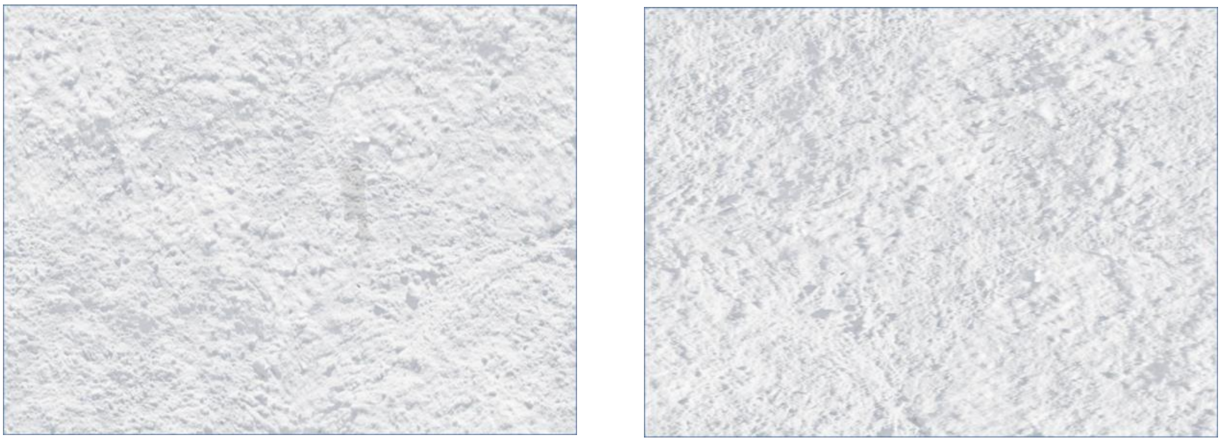

Figure 1. Example of critical stimuli present (left) and critical stimuli absent (right)

Figure 2. The three possible bodies in the critical stimulus conditions.

experimental blocks. The participants were required to signal their detection of the critical stimulus by pressing the response key on a keyboard (the spacebar); participants were to otherwise withhold from responding. Only responses made during the 2000ms of image presentation were recorded, responses made while the critical stimuli were present were recorded as correct detections (hits); responses while the critical stimuli were absent were recorded as errors of commission (false alarms). The images were followed by a blank inter stimulus interval (ISI) screen that was presented for 500ms to 1500ms. Participants in the points condition were shown their current running points during this period; points were calculated based on the participants' hits and false alarms; starting with a base of 100, points were increased by 10 for every correct detection and a decreased by 1 for every false positive. The points were weighted such that responding indiscriminately would, on average, neither increase nor decrease participants running score. Participants not in the points condition were not shown their running points during the ISI but were instead shown a message that instructed them simply to "please wait". The intermittent screen was followed by a fixation stimulus in which a crosshair was presented in the centre of the screen for 250ms on a white background. Each image, ISI, and fixation stimulus event took an average of 3000ms.

Before the main task began, participants were shown what they were to be looking for by example slides containing the target stimulus. The main task had a duration of approximately 22 minutes and featured 440 images presented at a rate of 20 events/minute.

3.2.2 Functional Near-Infrared Spectroscopy

For the functional near-infrared spectroscopy a Nonin® Model 7600 Cerebral Oximeter was used. The device consists of a 30cm x 18cm x 13cm monitor and a 4.3m trunk cable connecting to two Model 8000CA sensors as seen in figure 3; the sensors utilise light emitting diodes (LEDs) using three wavelengths in the 700 to 900 nanometre range to measure cerebral tissue oxygen saturation. The Nonin's dual-emitter sensor design allows for the compensation of surface artefacts and shallow tissue variation which may interfere with measurement accuracy as seen in figure 4. Sensor measurements of the subject's regional oxygen saturation (rSO_2) levels in both the left and right hemisphere as they engage in the specific tasks are recorded near instantly and can be monitored in real time on the device's display with a sampling rate 4 seconds apart. Participants' oxygen saturation levels before, during, and after the experimental session were transferred via Bluetooth wireless connection to a computer for later data analysis.

Figure 3. The EQUANOX™ Model 7600 Regional Oximeter System with Model 8000CA Sensors.

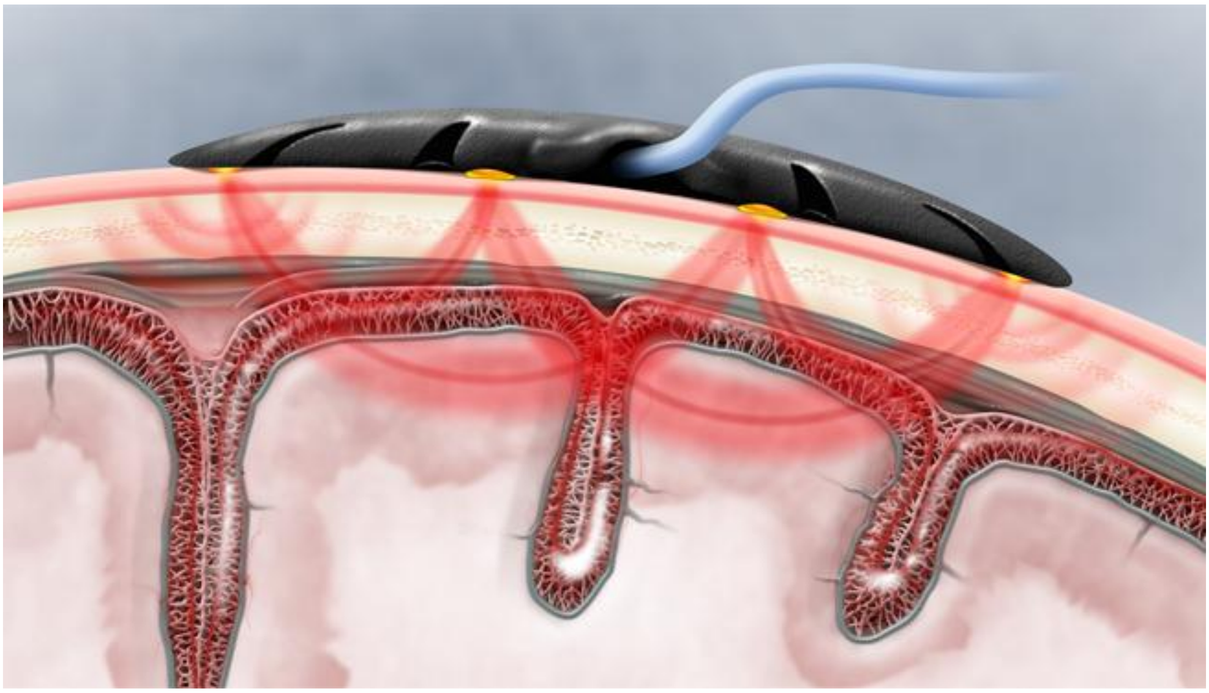

Figure 4. The functional near infrared spectroscopy sensor attached to the forehead emitting near infrared light with dual LEDs and receiving the reflected light with the two detectors. Photograph from http://www.nonintest.com/noninequanox/technology-overview/

3.2.3 *Dundee Stress State Questionnaire*

Participant's motivation was assessed through two sub-scales of the Dundee Stress State Questionnaire (Appendix A; Matthews, Joyner, Gilliland, Huggins, & Falconer, 1999; Matthews et al., 2002): Success Motivation (SM) and Intrinsic Motivation (IM). Both scales measured self-reported subjective states of motivation to perform the experimental task. The scales were administered two times during the course of the experiment; in the pre-task questionnaire immediately before the experimental task and in the post-task questionnaire immediately after. In both the pre and post-task questionnaires participants were required to quickly rate their agreement on a 5-point Likert scale (with 1 being 'not at all' and 5 being 'extremely') for each of the 15-items. The instructions emphasised that participants answer according to how they felt at that moment, not how they usually felt, in order to try and access their current state.

3.2.4 Utrecht Work Engagement Scale

Participant's task engagement was assessed through an adaptation of the Utrecht Work Engagement scale (UWES). The scale measured self-reported task-related states characterized by the sub-scales: Vigour (VI), Dedication (DE), and Absorption (AB). The scales were administered in the post-task questionnaire following the motivation scales. Participants were required to carefully rate their agreement on a 5-point Likert scale (with 1 being 'not at all' and 5 being 'extremely') for each of the 14-items

3.3 Procedure

The 40 participants were randomly assigned to one of four conditions resulting from the factorial combination of game mechanic assignment (narrative and control condition) and performance feedback (points and no-points control condition). All 40 participants were individually tested in the University's virtual reality lab which was a quiet, well-lit laboratory room with no external windows. Upon arrival participants were presented with an information sheet and a consent form which explained the aim of the project, the materials that were to be used, and the nature and duration of the participants' involvement. Once participants gave their informed consent they were asked to switch off any alarm or mobile device that they had on them to ensure that they were not distracted during the experiment. Participants were seated in front of a video display terminal (VDT) and were fitted with the fNIRS cerebral oximeter. The two fNIRS sensors were placed bilaterally on the participant's forehead (superior to the eyebrow and inferior to the hairline) with care taken so that sensors avoided being placed over hair, surface blemishes or sinus cavities which may have otherwise resulted in inaccurate readings. Once positioned in the appropriate place the sensors were secured to the head with an adjustable strap. Prior to the start of the experimental session a five minute baseline was conducted. During the baseline period participants were instructed to sit still and stare at the VDT in front of them (which was blank) and to maintain a state of relaxed wakefulness. The baseline period allowed for participants to become acclimated to the fNIRS procedure; cerebral oxygen saturation in the final 60 seconds of the baseline period provided the baseline index to be used in analysis (Ossowsk, Malinen & Helton, 2011; Stevenson, Russell & Helton, 2011).

After the baseline was taken, participants in the narrative condition were shown an 80 second video clip on the VDT. The video clip detailed the nature and danger of avalanches, organised rescue team's importance in rescuing buried avalanche victims, and a (fictional) ground penetrating radar used to detect the buried victims. Participants not in the narrative condition were not shown the video clip and went straight onto the pre-task questionnaire which participants completed through the computer terminal in front of them.

Following the completion of the pre-task questionnaire participants began the experimental task. Participants in narrative condition were given a brief outline of the scenario, which stated that an organised rescue team was responding to an avalanche disaster at Mt Hutt (a popular New Zealand high alpine ski field) and, as a member of that team, they had to find buried avalanche victims by examining the ground penetrating radar images. Those not in the narrative condition were simply told that they were looking for body shaped stimuli and were not given any additional background information.

Upon completion of the experimental tasks and questionnaires the fNIRS equipment was removed from the participant and they were debriefed about the purpose of the experiment.

4. Results

4.1 Performance

The performance metrics (proportion correct detections, proportion false alarms, positive predictive value, negative predictive value, d' and reaction time) were each analysed with a 2 (points versus no points) x 2 (narrative versus no narrative) x 4 (blocks of the task) split-plot analyses of variance (ANOVAs). For the block main effect and interactions, the primary interest was linear trend changes over blocks (representing differential learning gains) and therefore, the trend analysis results for these tests were conducted and reported, instead of the omnibus results. Only the significant results of these analyses (α = .05) and trends (α = .10) are reported; non-reported findings were therefore neither statistically significant nor indicative of a non-significant trend. The performance metrics for each block are presented in Appendix D.

4.1.1Correct Detections

For each individual for each block the proportion of correct detections was calculated [correct detection / (correct detections + misses)]. The overall rate of correct detections was moderately low (*M* = .686, *SE* =.024). There was a significant linear trend for blocks $F(1,36) = 5.561$, $p = .024$, $\eta_p^2 = .134$ with participants increasingly improving in performance with time spent on task, see figure 5. There was however, a significant linear trend between block, points, and narrative $F(1,36) = 5.561$, $p = .024$, $\eta_p^2 = .134$. This three way interaction appeared to result from a slightly higher increase in initial hit rate (block 1) in the story no-points condition (*M* = .70) relative to the other groups (*Ms* = .58 - .63).

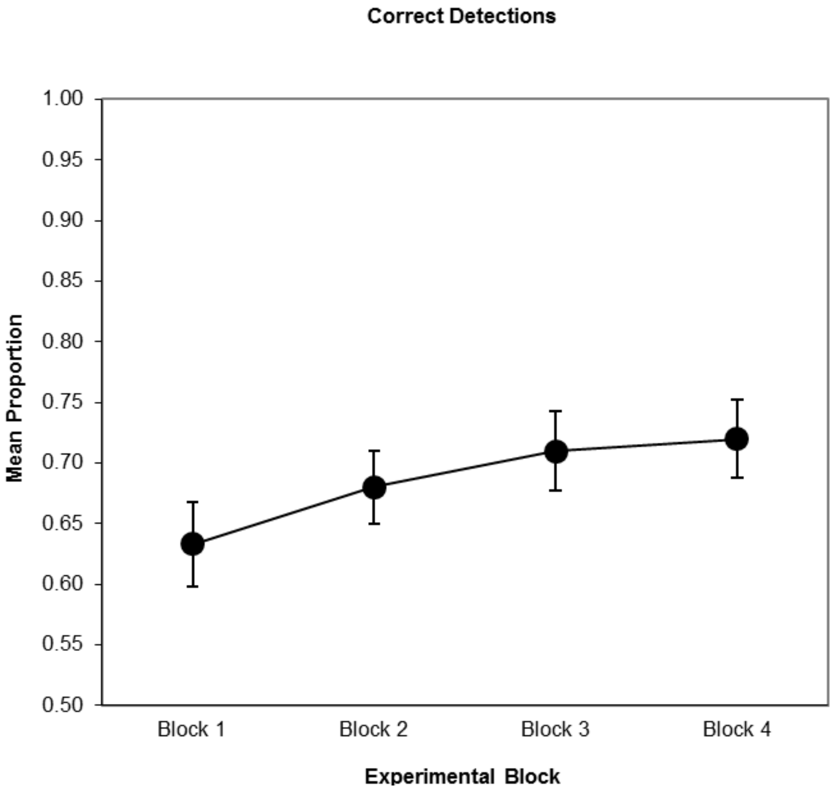

Figure 5. Mean proportion of correct detections over experimental block (error bars are standard error to the mean)

4.1.2 False Alarms

For each individual for each block the proportion of false alarms was calculated [false alarms / (false alarms + correct rejections)]. False alarm occurred relatively rarely with an overall rate of .046 (*SE* =.010). There was a significant linear trend for blocks $F(1,36) = 16.226$, $p < .001$, $\eta_p^2 = .311$ with the number of false alarms steadily declining with time spent on task, see figure 6. The points condition had a significant effect on performance $F(1,36) = 4.146$, $p = .049$, $\eta_p^2 = .103$ with participants in the points condition committing fewer false alarms ($M = .025$) than those in the no points condition ($M = .068$).

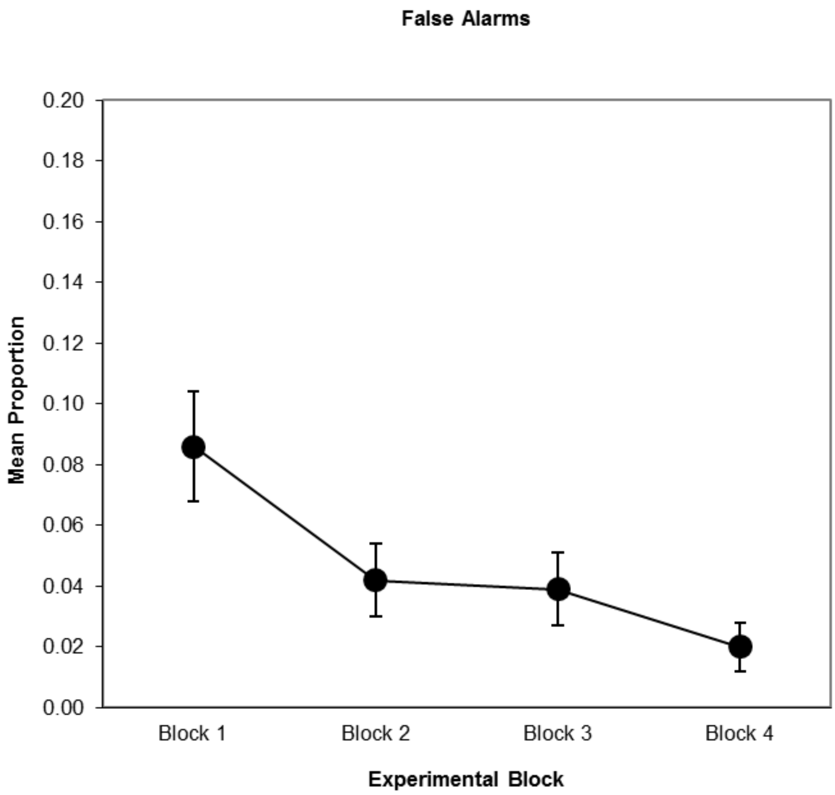

Figure 6. Mean proportion of false alarms over experimental block (error bars are standard error to the mean)

4.1.3 d'

d' is defined as the z-transformation of hits subtracted from the z-transformation of false alarms. *d'*, also known as the sensitivity index, is a statistic used in signal detection theory as an estimate of the signals strength. For each individual for each block *d'* was calculated. There was a significant linear trend for blocks $F(1,36) = 42.087$, $p < .001$, $\eta_p^2 = .539$, see figure 7.

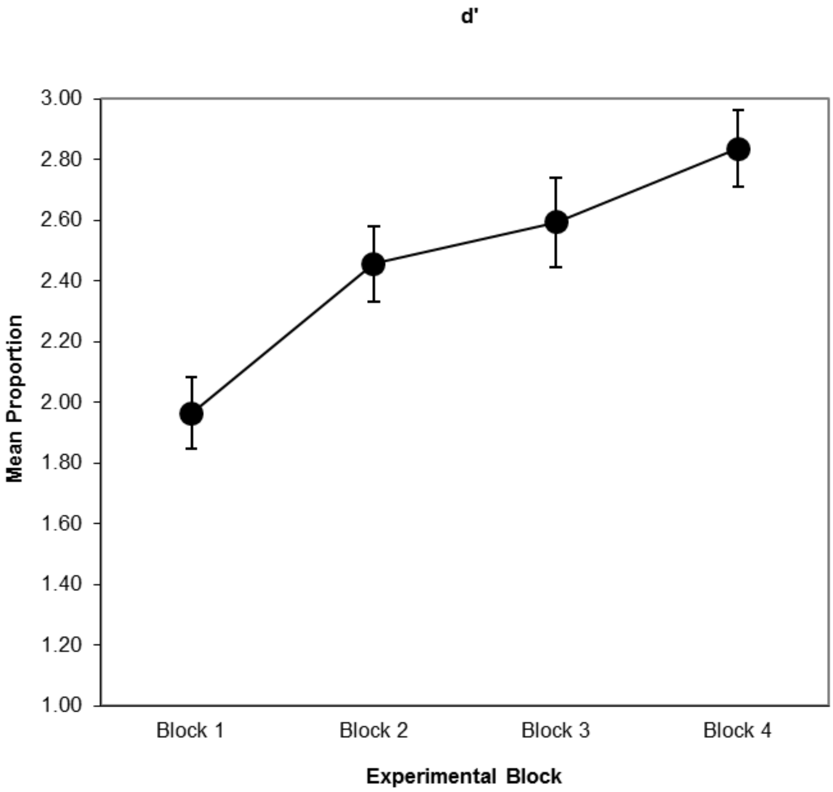

Figure 7. Mean d' over experimental block (error bars are standard error to the mean)

4.1.4 Positive Predictive Value

Positive predictive value (PPV) is a summary statistic defined as the number of true positives (correct detections) divided by the number of both true and false positives. PPV is a measure of precision and reflects the proportion of positive calls that were correctly diagnosed. For each individual for each block PPV was calculated. The overall PPV for participants was .752 (SE = .032). There was a significant linear trend for blocks $F(1,36)$ = 44.825, $p < .001$, η_p^2 = .555 with the positive predictive value increasing with time spent on task, see figure 8. The points condition had a significant effect on performance $F(1,36)$ = 5.191, p =.029, η_p^2 = .126 with participants in the points condition having a higher positive predictive value (M = .826) than those in the no points condition (M = .678).

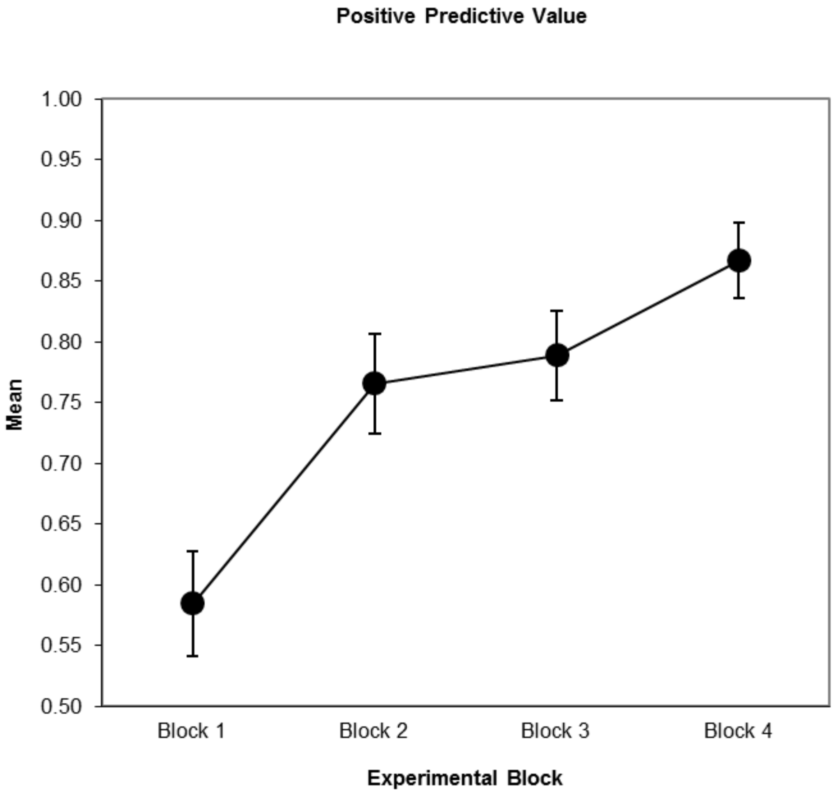

Figure 8. Mean positive predictive value over experimental block (error bars are standard error to the mean)

4.1.5 Negative Predictive Value

Negative predictive value (NPV) is a summary statistic defined as the number of true negatives (correct rejections) divided by the number of both true and false negatives. Like PPV, NPV is a measure of precision it reflects the proportion of negative calls that were correctly diagnosed. For each individual for each block NPV was calculated. The overall NPV was quite high (M = .968, SE = .002). There was a significant linear trend effect across the blocks $F(1,36) = 7.550$, $p = .009$, $\eta_p^2 = .173$ with participants' NPVs increasing with time spent on task, see figure 9.

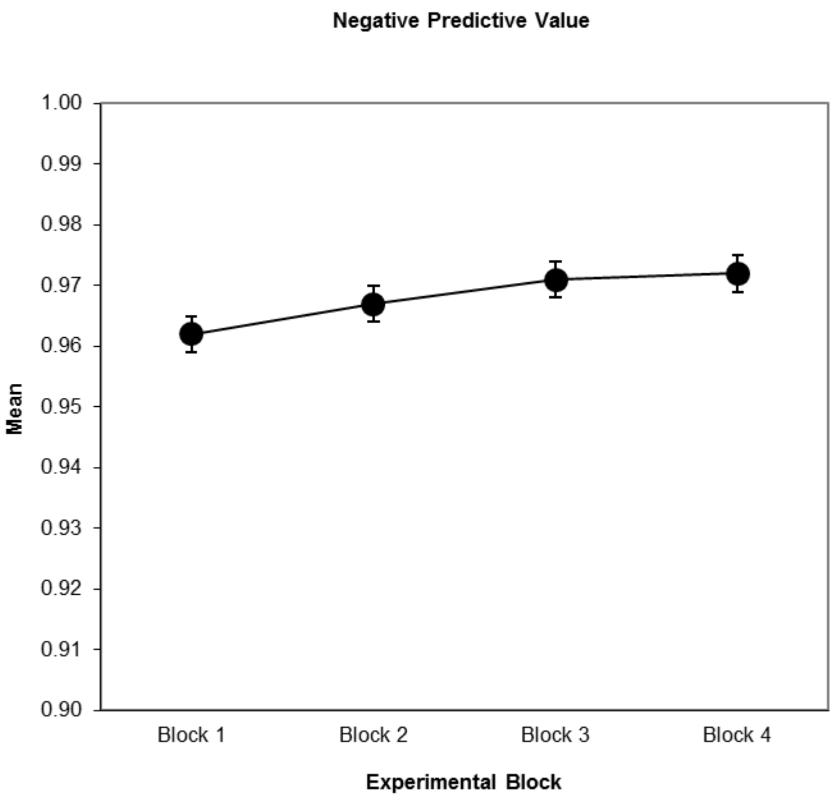

Figure 9. Mean negative predictive value over experimental block (error bars are standard error to the mean)

4.1.6 Reaction Times

Reaction times were defined as the median number of milliseconds since the onset of the stimulus that participants took to respond for each block. There was a significant linear trend for block $F(1,36) = 14.037$, $p = .001$, $\eta_p^2 = .281$ with participants' reaction times decreasing with time spent on task, see figure 10. The main effect for points was indicative of a trend, $F(1,36) = 3.212$, $p = .082$, $\eta_p^2 = .082$, with faster reaction times in the points condition (M = 1044 ms) relative to the no points condition (M = 1133 ms).

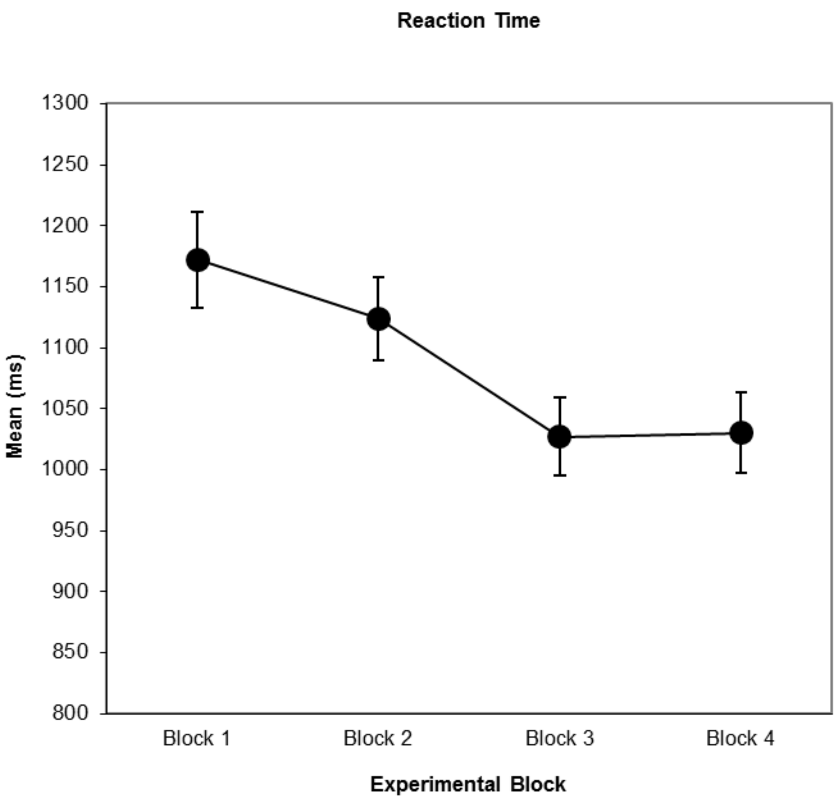

Figure 10. Mean reaction time over experimental block (error bars are standard error to the mean)

4.2 Physiology

As recommended in previous fNIRS studies (Yoshitani, Kawaguchi, Tatsumi, Kitaguchi & Furuya, 2002) a relative measure of regional oxygen (rSO_2) is used to analyse the data. This involves using the baseline taken prior to the experimental task as a benchmark to compare the percentage change in rSO_2. The cerebral rSO_2 levels for two participants were removed because of the fNIRS sensors slipping from position during the experiment. The relative rSO_2 levels were subjected to a 2 (hemisphere: left, right) x 2 (points verses no points) x 2 (narrative verses no narrative) x 4 (blocks) mixed ANOVA. The analysis revealed a significant linear trend for block $F(1,34) = 7.457$, $p = .010$, $\eta_p^2 = .180$. Additionally, there was also a close to significant main effect for story $F(1,34) = 4.016$, $p =$

.053, η_p^2 = .106 with participants in the narrative condition (*M* = 1.126, *SE* = .545) having higher rSO_2 levels relative to the baseline that participants in the no narrative condition (*M* = -.466, *SE* = .578).

4.3 Subjective Self-Reports

All items on the DSSQ and the UWES were measured all measured on the same scale (1 = Not at all, 5 = Extremely) therefore the un-standardised scores were used in analysis.

4.3.1 DSSQ

In the analysis the totals of the post-task questionnaires for both the Intrinsic Motivation scale (IM) and the Success Motivation scale (SM) of the post-task questionnaire were subtracted from the respective pre-task questionnaire totals to give an IM and SM change value. The change values were subjected to a 2 (narrative verses no narrative) x 2 (points verses no points) x 2 (scale: IM, SM) mixed ANOVA. The analysis revealed a significant linear trend for scale by points by story $F(1,35) = 4.578$, $p = .039$, see appendix G.

4.3.2 UWES

In the analysis the subjective ratings of Vigour (VI), Dedication (DE), and Absorption (AB) were used to examine participant engagement after the task. The data was analysed using a 2 (points verses no points) x 2 (narrative verses no narrative) x 3 (scale: VI, DE, AB) mixed ANOVA. The analysis revealed no significant results, $p > .05$.

4.4 Relationships between regional oxygenation and pre-task self-reported motivation and the performance metrics

In order to test specific predictions regarding the relationship between the regional oxygenation and pre-task self-reported motivation, two derived indices were calculated for each performance metric for each individual: the intercept and slope. Within each avalanche search task, lines of best fit using least squares estimation were calculated for each participant for both all six performance metrics. The four blocks within the avalanche search task were centred before calculating the lines of best fit. This was achieved by coding the four blocks sequentially as -1.5, -.5, .5, and 1.5 instead of 1, 2, 3, and 4 when computing the lines of best fit for each participant. Using this procedure, the intercept of the fitted line for a participant equals the mean over the entire four blocks of the task. The slope of the line indicates the linear trend or the change in magnitude of the performance metric of the participant over the four blocks (Keppel & Zedeck, 2001). In other words, the slope is indicative of the learning rate. The analysis enabled an examination of how average (collapsing across hemisphere) cerebral regional oxygenation scores for the four blocks and pre-task motivation related to both overall task performance (the intercept) and, perhaps more importantly, the learning rate or linear change in performance over the task (the slope) (see Helton et al., 2008; Helton & Warm, 2008; Langer et al., 2010). Prefrontal cerebral oxygenation and Pre-task motivation related to both overall task performance and learning rates. These results are presented in Appendix E.

5. Discussion

5.1 General Discussion

The current research investigated the effects of gamification of a diagnosis tasks in regards to increased performance through perceptual learning. An experiment was conducted where participants were split into conditions involving different game mechanics and completed an avalanche search task for 22 minutes. Participants' performance was evaluated through a range of performance metrics and their motivation and engagement were evaluated though self-report questionnaires and near-infrared spectroscopy.

Hypothesis 1 was supported by the performance data. Participants as a group clearly demonstrated perceptual learning as there were significant linear improvements in all six performance metrics. The proportion of correct detections, positive predictive value, negative predictive value, and *d'* all increased over experimental blocks. The proportion of false alarms and median response times both decreased over experimental blocks. These results are consistent with the literature on perceptual learning tasks where observers are expected become increasingly efficient at discriminating between stimuli (Watanabe, Nanez, & Sasaki, 2001; Watanabe et al., 2002) as opposed to what would be expected in vigilance tasks where performance characteristically declines over time (Davies & Parasuraman, 1982; Helton et al., 2005; Helton, Shaw, Warm, Matthews & Hancock, 2008).

Hypothesis 2 was supported by the findings with the fNIRS and self-report data. Congruent with previous research which found that subjects with greater initial brain activation in parietal and prefrontal regions were subsequently in learning the task (Mukai et al., 2007); Oxygenation response in the first block was indeed predictive of the learning rate (slope) for both proportion of correct detections and negative predictive value. Given that negative predictive value is a relative measure of correct rejections and misses and that proportion of correct detections is a relative measure of hits and misses, this is indicative that initial cerebral response being predictive of reductions in misses. Those individuals who appear to have higher levels of initial cognitive effort, increases frontal cerebral activation, appear to have a greater decline in missing critical stimuli. Intriguingly, self-reported intrinsic motivation was also predictive of performance, but appeared to be predictive of overall levels of performance not learning rates per se. Pre-task intrinsic motivation was correlated positively with overall positive predictive values and overall d' levels. This suggests consciously reportable task motivation and physiologically indexed cognitive or attentional effort are likely to be distinct. fNIRS measures of fontal activation may be a reasonable objective indicator of initial cognitive effort useful in perceptual learning studies, although this can be supplemented with self-report measures of task-motivation which tap separate processes.

Hypothesis 3 was only partially supported by the data. As mentioned above, in perceptual discriminatory tasks it has been found that higher brain activation in attention-related areas leads to improved performance. It was hypothesised that the game mechanics of points and narrative may result in an increase in participants' motivation which in turn would lead to improved and sustained attention.

Participants in the points condition were found having both significantly higher positive predictive value and making significantly less false alarms in the search task compared with participants in the no points condition. In addition, although not significant, reaction times tended to be faster for those in the points condition compared to those who were not.

However, in terms of motivation, points appeared to have a suppressant effect on narrative; across participants there was a decrement in self-reported intrinsic motivation but it was found that without points the narrative significantly reduces this drop. This may be due to the points distracting from the narrative's motivating ability or to perhaps, points trivializing the seriousness of an avalanche search (one of the possible downsides of gamificaiton in serious settings). The findings suggest that improvements in performance in the points condition was due to knowledge of results from the performance feedback rather than through increase motivation from satisfied psychological needs.

Looking at the prefrontal activation, the narrative manipulation did significantly influence cerebral oxygenation response. Regional oxygenation increased more in the narrative than the non-narrative condition. Oddly though, the narrative manipulation appeared to have no significant impact on performance, aside from an initial boost in the hit rate. The increase in oxygenation may have been due to increased activity in frontal executive system, perhaps, due to either the additional processing of the story elements or emotional processing due to the depressing statistics of avalanche searches, or both. Negative stimuli have been found to be distracting and attention demanding in previous studies (Helton & Head, 2012; Helton et al., 2011; Helton & Russell 2011; Ossowski et al., 2011). Therefore, the increase may or may not be due to increased cognitive effort with the task itself, but this should be explored in future studies. In summary, the mechanic of points lead to higher levels of performance in the experimental task while the mechanic of narrative lead to higher levels of intrinsic motivation and prefrontal oxygenation.

Contrary to the hypothesis 4 there were no significant differences for game mechanic condition on work engagement. This may be due to the experimental task being too transient in nature be able to meaningfully impact such a stable state like work engagement. Schaufeli et al. (2006) assert that work engagement is a relatively persistent and pervasive state that does not focus on specific behaviours or events.

5.2 Practical and Theoretical Implications

The findings from the current study have potential future applications for the work environment; the mechanics of points and narrative may be used as a means to improve performance and motivation respectively. The study's findings also provide much needed contributions to the literature surrounding gamification; the results of the experiment confirm what has been found in the industry, game mechanics can have real and significant effects on organisational outcomes. These findings may be helpful in building a theoretical model on the process behind gamification.

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 overall performance. Notably, the two different measures were more strongly related to different performance indexes; indicating that fNIRS is predictive in a different way than self-reported motivation. Self-reports are conscious, whereas fNIRS measured cerebral activity may be unconscious. This suggests that self-reported motivation is not the only kind of motivation or metric of engagement that is important in these kinds of tasks. This has significant real world implications especially in employment contexts where you are not always able to trust a person's self-report of motivation or work engagement. For instance, if a manager asked a person whether they were motivated to learn the job, the person is unlikely to say no in such a situation. fNIRS however, is not vulnerable to this type of deceit (without something like biofeedback training) and therefore has excellent potential as a supplementary or alternative measure to assess a person's state.

5.3 Limitations and Future Research

There were several limitations to this study. First, the study was conducted in a laboratory experiment context, as a result, the one off experimental task may not have been capable of meaningfully influencing a chronic state like work engagement (Schaufeli et al., 2006). Indeed, the UWES questionnaire that was used to measure work engagement had to be altered in the present study so as to be relevant to the context of the experiment; 3 of the 17 original items were omitted, such as "When I get up in the morning, I feel like going to work." which may have impacted its validity. In addition, while narrative did not enhance performance in the four experimental blocks statistically, there may have been other longer term effects, such as enhanced retention that were not evident in the current experiment. Future research could be directed at investigating gamification in a longitudinal study in order to properly investigate work engagement and explore any potential longer lasting effects.

Another possible limitation was that the nature of the experimental task may have been too emotionally evoking, especially for participants in the narrative condition where the direness and importance of the task were made more salient. This may have undermined any potential performance effect as negative stimuli have been found to both impair performance and distract attention (Helton & Head, 2012; Helton et al. 2011; Helton & Russell, 2011; Ossowski et al., 2011). Future research may wish to employ a less emotionally charged task as to avoid potentially contaminating the effect of narrative.

Finally, due to time and resource constraints, the current study had to limit its scope to only a small section of the incredibly broad area of gamification. There is great need for future studies into additional game mechanics and other aspects of game design in order to gain a more complete understanding of gamification.

5.4 Concluding Statement

The study revealed that the game mechanics of points and narrative can have a significant impact on both perceptual learning performance and intrinsic motivation respectively, these results support the findings in the industry that demonstrate the ability of gamification to impact organisational outcomes. The study also revealed that initial oxygen response and self-reported pre-task motivation has differential predictiveness, and they were more strongly related to different performance indexes. This finding presents significant real world applications for objectively measuring motivation.

The science of gamification is a complex topic that has only just started to be explored. As game mechanics and their underlying mechanisms become better understood they have great potential to be successfully applied to the business environment.

6. References

Alexander, P., Ryan, R., & Deci, E. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. Contemporary Educational Psychology, 25, 1.

Ayaz H., Shewokis P. A., Bunce S., Izzetoglu K., Willems, B., Onaral, B. (2012). Optical Brain Monitoring for Operator Training and Mental Workload Assessment. *NeuroImage*, 59(1), 36-47. doi: 10.1016/j.neuroimage.2011.06.023.

Bransford, J., Brown, A., Cocking, R., (2000). How people learn. National Research Council, National Academy Press.

Bourque, A. (2012) *Not All Fun and Games: One Call Center's Performance Success through Gamification.* Social Media Today. Retrieved from http://socialmediatoday.com/socialmktgfella/483538/call-centers-performance-success-through-gamification

Burke, M., & Hiltbrand, T. (2011). How Gamification Will Change Business Intelligence. Business Intelligence Journal, 16(2), 8-16. Data Warehousing Institute. Retrieved from http://proxy2.hec.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=61965733&lang=fr&site=bsi-live

Caplan, S., D. Williams & N. Yee. (2009). Problematic internet use and psychosocial well-being among *MMO players*. Computers in Human Behavior. 25(6) p. 1312-1319.

Castel, A.D., Pratt, J., Drummond, E. (2005). The effects of action video game experience on the time course of inhibition of return and the efficiency of visual search. Ada Psychol (Amst), 119(2):217–230. doi:10.1016/j.actpsy.2005.02.004

Csíkszentmihályi, M. (1990). Flow: The Psychology of Optimal Experience. New York: Harper and Row.

Csikszentmihalyi, M., & Rathunde, K. (1993). The measurement of flow in everyday life: Towards a theory of emergent motivation. In J. E. Jacobs (Ed.), Nebraska symposium on motivation, Vol.40: Developmental perspectives on motivation (pp. 57-98). Lincoln: University of Nebraska Press.

Curtiss, M. (2011). "Why Games Work and the Science of Learning" retrieved from http://www.goodgamesbydesign.com/Files/WhyGamesWork_TheScienceOfLearning_CMurCMu_2011.pdf

Danner, F. W., & Lonky, E. (1981). A cognitive-developmental approach to the effects of rewards on intrinsic motivation. Child Development, 52, 1043–1052.

Davies, D.R., & Parasuraman, R. (1982). *The psychology of vigilance.* Jondon: Academic Press.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology, 18,* 105–115.

Deci, E. L., & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behavior. New York: Plenum

Deterding, S., Khaled, R., Nacke, L.E., Dixon, D. (2011). Gamification: Toward a Definition. In CHI 2011 Gamification Workshop Proceedings, Vancouver, BC, Canada.

Fairclough, S. H., Venables, L., & Tattersall, A. J. (2005). The influence of task demand and learning on the psychophysiological response. *International Journal of Psychophysiology, 56*, 171-184.

Fisher, C. D. (1978). The effects of personal control, competence, and extrinsic reward systems on intrinsic motivation. *Organizational Behavior and Human Performance, 21*, 273–288

Gartner Group (2011) Gartner says by 2015, more than 50 percent of organizations that manage innovation processes will gamify those processes. Gartner Enterprise

Architecture Summit Press release retrieved from http://www.gartner.com/newsroom/id/1629214,

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

Google trends "Gamification at Google Trends". Retrieved from http://www.google.com/trends/explore#q=gamification 2013-01-31.

Green, C.S., & Bavelier, D. (2006) Effect of action video games on the spatial distribution of visuospatial attention. Journal of Experimental Psychology: Human Perception and Performance, 32(6):1465–1478. doi:10.1037/0096-1523.32.6.1465

Greenfield, P.M., DeWinstanley, P., Kilpatrick, H., Kaye, D. (1994). Action video games and informal education: effects on strategies for dividing visual attention. Journal of Applied Developmental Psychology,15:105–123. doi:10.1016/0193-3973(94)90008-6

Harter, J.K., Schmidt, F.L., Killham, E.A., & Agrawal, S. (2009). Gallup Q12® meta-analysis: the relationship between engagement at work and organizational outcomes. Omaha, NE: Gallup.

Heckhausen, H. (1967). *The Anatomy of Achievement Motivation.* Academic Press, New York

Helton, W.S., Hollander, T.D., Warm, J.S., Matthews, G., Dember, W.N., Wallart, M., Deauchamp, G., Parasuraman, R., & Hancock, P.A. (2005). Signal regularity and the mindlessness model of vigilance. *British Journal of Psychology, 96,* 249-261. 39

Helton, W.S., Shaw, T., Warm, J.S., Matthews, G., & Hancock, P.A. (2008). Effects of warned and unwarned demand transitions on vigilance performance and stress. *Anxiety, Stress and Coping, 21,* 173-184.

Helton, W. S., & Russell, P. N. (2011). The effects of arousing negative and neutral picture stimuli on target detection in a vigilance task. *Human Factors, 53,* 132-141.

Helton, W. S., & Warm, J.S. (2008). Signal salience and the mindlessness theory of vigilance. *Acta Psychologica, 129,* 18-25. Helton, W.S., Shaw, T., Warm, J.S., Matthews, G., & Hancock, P.A. (2008). Effects of warned and unwarned demand transitions on vigilance performance and stress. *Anxiety, Stress and Coping, 21,* 173-184.

Helton, W.S., Dorahy, M.J., & Russell, P.N. (2011). Dissociative tendencies and right hemisphere processing load: effects on vigilance performance. *Consciousness and Cognition, 20,* 696-702

Helton, W.S., Head, J., & Kemp, S. (2011). Natural disaster induced cognitive disruption: impacts on action slips. *Consciousness and Cognition, 20,* 1732-1737.

Helton, W.S., & Head, J. (2012). Earthquakes on the mind: implications of disasters for human performance. *Human Factors.*

Herger, M. (2011). "Gamification Facts & Figures". Retrieved from http://enterprise-gamification.com/index.php/en/facts

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

Herzog, M.H. & Fahle, M. (1998). Modeling perceptual learning: Difficulties and how they can be overcome. Biological Cybernetics, 78(2), 107-117. doi:10.1007/s004220050418

Huppert, T. J., Hoge, R.D., Diamond, S.G., Franceschini, M.A., Boas, D.A., 2006. A temporal comparison of BOLD, ASL, and NIRS hemodynamic responses to motor stimuli in adult humans. NeuroImage 29, 368-382.

Hussain T. S., Roberts B., Bowers, C. Cannon-Bowers, J. A. Menaker, E. S. Coleman, S. L. Murphy, C. Pounds, K. Koenig, A. Wainess, R., Lee, J. “Designing and developing effective training games for the US Navy," in Proceedings of the Interservice/Industry Training, Simulation, and Education Conference, 2009

Ishii Y., Ogata, H., Takano, H., Ohnishi, H., Toshiharu Mukai, T., Yagi, T. (2008). Study on mental stress using near-infrared spectroscopy, electroencephalography, and peripheral arterial tonometry Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE, pp. 4992 – 4995

Jacobs, S. (2011): Blog: Serious Play Conference 2011: Microsoft's 'Productivity Games' , retrieved from http://www.gamasutra.com/view/news/36824/Serious_Play_Conference_2011_Microsofts_Productivity_Games.php#.UPEwRGNFunA

Keppel, G. & Zedeck, S. (2001). *Data analysis for research designs.* New York: W.H. Freeman & Co.

Koepp, M. J., Gunn, R. N., Lawrence, A. D., Cunningham, V. J., Dagher, A., Jones, T., Brooks, D. J., Bench C. J., Grasby, P. M. (1998) Evidence for striatal dopamine release during a video game. *Nature*, 393(6682):266-267.

Langner R, Willmes K, Chatterjee A, Eickhoff SB, Sturm W (2010) Energetic effects of stimulus intensity on prolonged simple reaction time performance. Psychological Research, 74, 499-512.

Mukai I, Kim D, Fukunaga M, Japee S, Marrett S, Ungerleider LG. Activations in visual and attention-related areas predict and correlate with the degree of perceptual learning. Journal of Neuroscience. 2007;27:11401–11411.

Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology, 52*, 397-422.

Matthews G., Campbell S. E., Falconer S. (2001). Assessment of motivational states in performance environments. In *Proceedings of the Human Factors and Ergonomics*

Society 45th Annual Meeting (pp. 906–910). Santa Monica, CA: Human Factors and Ergonomics Society.

Matthews, G., Campbell, S.E., Falconer, S., Joyner, L.A., Huggins, J., Gilliand, K., Grier, R., & Warm, J.S. (2002). Fundamental dimensions of subjective state in performance settings: Task engagement, distress, and worry. *Emotion*, *2,* 315-340.

Matthews, G., Joyner, L., Gilliland, K., Campbell, S. E., Huggins, J., &Falconer, S. (1999). Validation of a comprehensive stress state questionnaire: Towards a state "Big Three"? In I. Mervielde, I. J. Deary, F.De Fruyt, & F. Ostendorf (Eds.), Personality psychology in Europe (Vol.7, pp. 335–350). Tilburg, the Netherlands: Tilburg University Press

McClelland, D. C., J. W. Atkinson, R. A. Clark, E. L. Lowell. 1953. *The Achievement Motive*, Century psychology series. Appleton-Century-Crofts, East Norwalk, CT.

McGonigal, J., (2011). *Reality Is Broken: Why Games Make Us Better and How They Can Change the World*, Penguin Press HC.

Medland, B. (2013) Video game industry is big business. The Brock Press retrieved from http://www.brockpress.com/news/business/video-game-industry-is-big-business-1.2978741#.UQmynWNFunA

Nike plus, Retrieved from http://nikeplus.nike.com/plus/

Ossowski, U., Malinen, S. Helton, W.S. (2011). The effects of emotional stimuli on target detection: Indirect and direct resource costs. *Consciousness and Cognition, 20,* 1649–1658.

Pink, D. (2009). Dive – The Surprising Truth About What Motivates Us. Penguin Group.

Parasuraman, R. & Wickens, C.D. (2008). Humans: Still vital after all these years of automation. *Human Factors, 50,* 511-520.

Podsakoff, P.M.,& Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management, 12*(2): 531-544.

Plichta, M.M., Herrmann, M.J., Baehne, C.G., Ehlis, A.C., Richter, M.M., Pauli, P., Fallgatter, A.J. (2007). Event-related functional near-infrared spectroscopy (fNIRS) based on craniocerebral correlations: reproducibility of activation? *Human Brain Mapping 28,* 733-741.

Porter, L. W., & Lawler, E. E. (1968). Managerial Attitudes and Performance. Homewood, IL: Richard D. Irwin, Inc.

Reeves, B. & Read, J.L. (2009). *Total engagement: using games and virtual worlds to change the way people work and businesses compete*. Harvard Business School Press: Boston

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology*, *43*, 450-461.

Ryan, R. M., Rigby, C. S., & Przybylski, A. K. (2006). Motivation pull of video games: A Self-determination theory approach. *Motivation and Emotion*, 30, 347-365

Saatchi & Saatchi (2011). Engagement unleashed: Gamification for business, brands, and loyalty. Retrieved from www.slideshare.net/Saatchi_S/gamification-study.

Schaufeli, W.B., Salanova, M., González-Romá, V., & Bakker, A. B. (2002). The measurement of Engagement and burnout: A confirmative analytic approach. *Journal of Happiness Studies*, 3, 71-92

Silverman, R. (2011). "Latest Game Theory: Mixing Work and Play - Companies Adopt gaming Techniques to Motivate Employees". Wallstreet Journal. Retrieved from http://online.wsj.com/article/SB10001424052970204294504576615371783795248.html

Shronfeld, E. (2010). SCVNGR's Secret Game Mechanics Playdeck. Retrived from http://techcrunch.com/2010/08/25/scvngr-game-mechanics/

Steers, R.M., Mowday, R.T., & Shapiro, D.L. 2004 The future of work motivation theory. Academy of Management Review, 25 (3): 379-387.

Stevenson, H., Russell, P.N., & Helton, W.S. (2011). Search asymmetry, sustained attention, and response inhibition. *Brain and Cognition, 77*, 215-222.

Thorndike, E. (1932). *The Fundamentals of Learning*. New York: Teachers College Press.

Williams, K., Williams C. (2010) “Five key ingredients for improving student motivation”. Research in Higher Education Journal.

Watanabe, T., Nanez, J. E. & Sasaki, Y. (2001). Perceptual learning without perception. Nature 413, 844–848

Watanabe, T., Nanez, J. E., Sr., Koyama, S., Mukai, I., Liederman, J., & Sasaki, Y. (2002). Greater plasticity in lower-level than higher-level visual motion processing in a passive perceptual learning task. Nature Neuroscience, 5, 1003–1009.

Willis, J.A. (2010), 'Neurologist Makes the Case for the Video Game Model as a Learning Tool', Edutopia, Retrieved from http://www.edutopia.org/blog/video-games-learning-student-engagement-judy-willis.

Wright, B.E. (2001). Public sector work motivation: Review of current literature and a revised conceptual model. *Journal of Public Administration Research and Theory*, *11*(4), 559-586.

Xu, Y., 2012. *Literature Review on Web Application Gamification and Analytics*. Honolulu: University of Hawaii. Retrieved from: https://csdl-techreports.googlecode.com/svn/trunk/techreports/2011/11-05/11-05.pdf.

Zichermann, G., Cunningham, C. (2011)."Preface". Gamification *by Design: Implementing Game Mechanics in Web and Mobile Apps* (1st ed.). Sebastopol, California:O'Reilly Media. p. ix. ISBN 1449315399. Retrieved 2012-11-25. "Gamification may be a new term"

Appendix A

UC UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Demographics

General Instructions

This questionnaire is concerned with your feelings and thoughts at the moment. Please answer **every** question, even if you find it difficult. Answer, as honestly as you can, what is true of **you**. Please do not choose a reply just because it seems like the 'right thing to say'. Your answers will be kept entirely confidential. Also, be sure to answer according to how you feel **AT THE MOMENT**. Don't just put down how you usually feel. You should try and work quite quickly: there is no need to think very hard about the answers. The first answer you think of is usually best.

Before you start, please provide some general information about yourself.

Sex

- ○ Male
- ○ Female

What year were you born?

[▾]

Please answer some questions about your attitude to the task you are about to do. Rate your agreement with the following statements by selecting one of the following answers:

	Not at	A little bit	Somewhat	Very much	Extremely
I expect the content of the task will be interesting	○	○	○	○	○
The only reason to do the task is to get an external reward	○	○	○	○	○
I would rather spend the time doing the task on something else	○	○	○	○	○
I am concerned about not doing as well as I can	○	○	○	○	○
I want to perform better than most people do	○	○	○	○	○
I will become fed up with the task	○	○	○	○	○
I am eager to do well	○	○	○	○	○
I would be disappointed if I failed to do well on the task	○	○	○	○	○
I am committed to attaining my performance goals	○	○	○	○	○
Doing the task is worthwhile	○	○	○	○	○
I expect to find the task boring	○	○	○	○	○
I feel apathetic about my performance	○	○	○	○	○
I want to succeed on the task	○	○	○	○	○
The task will bring out my competitive drives	○	○	○	○	○
I am motivated to do the task	○	○	○	○	○

Appendix B

Please answer the following questions about your attitude to the task you have just done. Rate you agreement with the following statements by selecting one of the following answers:

	Not at all	A little bit	Somewhat	Very much	Extremely
The content of the task was interesting	○	○	○	○	○
The only reason to do the task is to get an external reward (e.g. payment)	○	○	○	○	○
I would rather have spent the time doing the task on something else	○	○	○	○	○
I was concerned about not doing as well as I can	○	○	○	○	○
I wanted to perform better than most people do	○	○	○	○	○
I became fed up with the task	○	○	○	○	○
I was eager to do well	○	○	○	○	○
I would be disappointed if I failed to do well on this task	○	○	○	○	○
I was committed to attaining my performance goals	○	○	○	○	○
Doing the task was worthwhile	○	○	○	○	○
I found the task boring	○	○	○	○	○
I felt apathetic about my performance	○	○	○	○	○
I wanted to succeed on the task	○	○	○	○	○
The task brought out my competitive drives	○	○	○	○	○
I was motivated to do the task	○	○	○	○	○

The following 14 statements are about how you felt during the task. Please read each statemetent carefully and dcide how much you agree with it.

	Not at	A little bit	Somewhat	Very much	Extremely
During the task, I felt bursting with energy	○	○	○	○	○
I found the task that I did full of meaning and purpose	○	○	○	○	○
During the task, I felt strong and vigorous	○	○	○	○	○
I was enthusiastic about the task	○	○	○	○	○
During the task I forgot everything else around me	○	○	○	○	○
The task inspired me	○	○	○	○	○
I am proud of the task I did	○	○	○	○	○
I was immersed in the task	○	○	○	○	○
I could continue doing the task for very long periods at a time	○	○	○	○	○
To me, the task was challenging	○	○	○	○	○
I got carried away when I was completing the task	○	○	○	○	○
During the task, I was very resilient, mentally	○	○	○	○	○
It was difficult to detach myself from the task	○	○	○	○	○
During the task, I always persevered, even when things were not going well	○	○	○	○	○

Debriefing

The effect of gamification on performance and engagement

Thank you for participating in this research project. The purpose of the research is to see if gamification (the use of game mechanics and game thinking to enhance non-game contexts) has an effect on performance and engagement. Specifically whether points and narrative increase engagement and performance and how they work in combination.

To test the effect of the game mechanics of points and narrative half of the experiments where done using them and half were done without using them. Participants completed the experiment both with/without the points mechanic and with/without the narrative mechanic. Your performance was determined by how fast and accurately you were and your engagement was determined by the post experiment questionnaire and by blood oxygen levels in the brain.

It is predicted that the game mechanics of points and narrative will improve the performance of the task and will increase the level of engagement. The implication of this is that future jobs may have game mechanics built into the job tasks.

For any additional information you may contact Michael Ong at michael.ong@pg.canterbury.ac.nz.

Appendix C

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Department of Psychology

Michael Ong

Michael.ong@pg.canterbury.ac.nz

22 June 2012

CONSENT FORM

The effect of gamification on performance and engagement

I have read and understood the description of the above-named project. On this basis I agree to participate as a subject in the project, and I consent to publication of the results of the project with the understanding that anonymity will be preserved.

I understand also that I may at any time withdraw from the project, including withdrawal of any information I have provided.

I note that the project has been reviewed ***and approved*** by the University of Canterbury Human Ethics Committee.

Signature:

Date:

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Department of Psychology

INFORMATION

You are invited to participate as a subject in the research project 'Gamification and its effect on performance and engagement'.

The aim of this project is to understand the relationship between game mechanics and work outcomes.

Your involvement in this project will be to complete a questionnaire and a computer based visual scanning task which will take approximately 30 minutes. During the experiment blood oxygen levels will be monitored via near infrared sensors attached to your forehead. You have the right to withdraw from the project at any time, including withdrawal of any information provided without penalty.

As a follow-up to this investigation, you will be asked to answer a short questionnaire about the task.

The results of the project may be published, but you are assured of the complete confidentiality of data gathered in this investigation: the identity of participants will not be made public without their consent. To ensure anonymity and confidentiality, the experiment and questionnaire is computer based so your personal information such as your name will not be used. Additionally since remuneration is not done through prize draw you do not have to leave any contact information.

By participating in this research project you will be compensated with a $15 Westfield voucher.

The project is being carried out as a requirement for Master of Science in Applied Psychology by Michael Ong under the supervision of Associate Professor William Helton, who can be contacted at +64 3 364 2998, ext. 7999. He will be pleased to discuss any concerns you may have about participation in the project.

The project has been reviewed ***and approved*** by the University of Canterbury Human Ethics Committee.

Appendix D

Table 1. Mean performance over the task blocks (standard errors of the mean).

	Block 1	Block 2	Block 3	Block 4	Linear Trend		
					F	*p*	η_p^2
Correct Detections	.633	.680	.710	.720	5.56	.024	.134
	(.035)	*(.030)*	*(.033)*	*(.032)*			
False Alarms	.086	.042	.039	.020	16.27	.000	.311
	(.018)	*(.012)*	*(.012)*	*(.008)*			
Positive Predictive Value	.585	.766	.789	.867	44.83	.000	.555
	(.043)	*(.041)*	*(.037)*	*(.031)*			
Negative Predictive Value	.962	.967	.971	.972	7.55	.009	.173
	(.003)	*(.003)*	*(.003)*	*(.003)*			
d'	1.965	2.457	2.594	2.836	42.09	.000	.539
	(.117)	*(.123)*	*(.147)*	*(.126)*			
Reaction Time	1172	1124	1027	1030	14.04	.001	.281
	(39)	*(34)*	*(32)*	*(33)*			

Appendix E

Table 2. The relationship between cerebral oxygenation and pre-task motivation and the performance metrics (correlation coefficients).

	Correct Detections		False Alarms		PPV		NPV		d'		Reaction Time	
	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope
rSO2 Block 1	0.19	**<u>0.37</u>**	0.00	0.03	-0.04	0.04	0.15	**<u>0.37</u>**	0.08	0.29	0.07	-0.23
rSO2 Block 2	0.19	**<u>0.36</u>**	-0.12	0.07	0.07	-0.07	0.18	**<u>0.34</u>**	0.18	0.21	-0.07	-0.11
rSO2 Block 3	0.15	**<u>0.32</u>**	-0.17	0.13	0.17	-0.15	0.16	0.29	0.23	0.14	-0.12	-0.06
rSO2 Block 4	0.19	0.25	-0.17	0.11	0.17	-0.15	0.20	0.22	0.25	0.10	-0.12	-0.03
Success Motivation	-0.16	-0.14	-0.05	-0.01	0.05	-0.11	-0.17	-0.11	-0.07	-0.19	-0.01	0.05
Intrinsic Motivation	0.20	0.01	-0.29	0.19	**<u>0.39</u>**	0.07	0.24	-0.01	**<u>0.38</u>**	0.04	-0.19	0.22

Note: Significant correlations ($p < .05$) are bold and underlined

Appendix F

Tests of within-subject effects for change in effects for SM and IM change values

8. Story * Points * scale

Measure:MEASURE_1

Story	Points	scale	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
NoStory	NoPoints	1	.302	.236	-.178	.782
		2	-.714	.184	-1.087	-.341
	Points	1	.086	.224	-.370	.541
		2	-.471	.174	-.825	-.117
Story	NoPoints	1	.114	.224	-.341	.570
		2	-.271	.174	-.625	.083
	Points	1	.286	.224	-.170	.741
		2	-.657	.174	-1.011	-.303

Tests of Within-Subjects Effects for SM and IM change values

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
scale	Sphericity Assumed	10.240	1	10.240	37.352	.000	.516
	Greenhouse-Geisser	10.240	1.000	10.240	37.352	.000	.516
	Huynh-Feldt	10.240	1.000	10.240	37.352	.000	.516
	Lower-bound	10.240	1.000	10.240	37.352	.000	.516
scale * Story	Sphericity Assumed	.073	1	.073	.265	.610	.008
	Greenhouse-Geisser	.073	1.000	.073	.265	.610	.008
	Huynh-Feldt	.073	1.000	.073	.265	.610	.008
	Lower-bound	.073	1.000	.073	.265	.610	.008
scale * Points	Sphericity Assumed	.012	1	.012	.043	.837	.001
	Greenhouse-Geisser	.012	1.000	.012	.043	.837	.001
	Huynh-Feldt	.012	1.000	.012	.043	.837	.001
	Lower-bound	.012	1.000	.012	.043	.837	.001
scale * Story * Points	Sphericity Assumed	1.255	1	1.255	4.578	.039	.116
	Greenhouse-Geisser	1.255	1.000	1.255	4.578	.039	.116
	Huynh-Feldt	1.255	1.000	1.255	4.578	.039	.116
	Lower-bound	1.255	1.000	1.255	4.578	.039	.116
Error(scale)	Sphericity Assumed	9.595	35	.274			
	Greenhouse-Geisser	9.595	35.000	.274			
	Huynh-Feldt	9.595	35.000	.274			
	Lower-bound	9.595	35.000	.274			