Not Doing But Thinking: The Role of Challenge in the Gaming Experience

Anna L Cox
UCL Interaction Centre
University College London
anna.cox@ucl.ac.uk

Paul Cairns
Dept of Computer Science
University of York
paul.cairns@york.ac.uk

Pari Shah
Psychology & Language Sciences
University College London
zcjtbb4@ucl.ac.uk

Michael Carroll
Dept of Computer Science
University of York
mjpc@cs.york.ac.uk

ABSTRACT
Previous research into the experience of videogames has shown the importance of the role of challenge in producing a good experience. However, defining exactly which challenges are important and which aspects of gaming experience are affected is largely under-explored. In this paper, we investigate if altering the level of challenge in a videogame influences people's experience of immersion. Our first study demonstrates that simply increasing the physical demands of the game by requiring gamers to interact more with the game does not result in increased immersion. In a further two studies, we use time pressure to make games more physically *and* cognitively challenging. We find that the addition of time pressure increases immersion as predicted. We argue that the level of challenge experienced is an interaction between the level of expertise of the gamer and the cognitive challenge encompassed within the game.

Author Keywords
Games; immersion; challenge.

ACM Classification Keywords
K.8.0 Personal Computing: General - Games

General Terms
Experimentation

INTRODUCTION
Videogames are possibly the most successful application of computing by almost any measure. For instance, the games industry now exceeds cinema in terms of sales [5] and eight out of ten homes in the UK own a next-generation videogame console [29]. What underlies this success is undoubtedly the experiences that people have with games, but what exactly those experiences are and how they influence people is only just beginning to be understood.

CHI'12, May 5–10, 2012, Austin, Texas, USA.

There are currently several approaches to defining and measuring gaming experience (GX) and whilst these differ in emphasis and detail, there are general overlaps in how gaming experience is understood. At the most basic level, GX arises from engagement with games and it is useful to both quantify that engagement [1] and to understand the components of a game that lead to engagement [4]. There are also concerns for specific aspects of gaming experience such as immersion, the sense of being cognitively involved in the game to the exclusion of the world around you [12, 15], and presence, the sense of being actually present in the world of the game as if it were real [26]. Both of these concepts are included in more general approaches to GX that attempt to cover all the major components of experiences that may arise from playing videogames not just immersion and presence but also enjoyment, flow, emotions and so on [20]. These more general descriptions of GX even extend to include wider, more complex aspects of gaming such as narrative and social relationships [21, 9].

These descriptions of GX provide frameworks in which to understand how different aspects of games lead to specific experiences. Research in this area covers diverse features from the controllers used to play [26] and the scenarios used to motivate the game [22] to the Quality of Service provided in networked games [30] and even the reviews games receive [18].

Underpinning the study of GX is the need for effective metrics. Those descriptions of GX mentioned above are operationalised as questionnaires that allow for the quantification of GX. Other richer, more complex measures are also possible such as the use of patterns to classify behaviour of players [10] or the use of log or telemetry data [13].

Whilst much is being done, there are still many fundamental questions that remain. The aim of this paper is to address one such question by considering the role of challenge in GX.

Immersion

In the work reported here, we are specifically concerned with the concept of immersion. This is the sense of being absorbed in a game to the exclusion of all else outside of the game. The experience of immersion in videogames is commonly reported by players, reviewers and designers of games [3] and, in various guises, is common to many of the theories of GX. Typical consequences of immersion are a sense of dissociation from the real world, a loss of the sense of time and a high degree of cognitive and emotional involvement in the gameplay and its outcome [15].

Whilst immersion is recognized as important in GX, the field of positive psychology is dominated by the theory of Flow [8]. Flow is an extreme experience combining elements of skill, progress, challenge and focus to produce an optimal experience. It is best characterized as the sporting experience of being "in the zone" but can also be experienced in factory work and playing chess. Within videogames, flow is considered to be an important component of GX [6, 7]. Others have tailored this theory specifically for gaming with the concept of GameFlow [27] but whereas flow is the experience had by players, GameFlow is the features of a game that are proposed to generally lead to flow experiences. As such, flow and GameFlow both relate to optimal experiences. However, when flow is formulated in the context of GX and particularly measuring GX, flow is actually operationalised to be a graded experience where people can grade their level of flow e.g. [20]. This makes little sense – an athlete cannot be "a bit in the zone" and gamers cannot be a bit in flow.

In contrast, immersion is acknowledged as a graded experience ranging from engagement, through engrossment to total immersion [3]. It is only total immersion that can be viewed as being analogous to flow [25]. When playing a videogame, many people are having much more prosaic experiences than those conceived by the notion of flow. When in flow, by definition, one is not monitoring other activities in the environment. But when immersed in a videogame, other things can be going on that the player is attending to. For instance, in mobile games, the player might also be waiting for a bus and thus actively monitoring the environment for the arrival of a bus, checking to see if it has the required route number as well as applying focused attention and engagement in playing a game on their phone. The player is somewhat immersed but definitely not in flow.

Challenge

Across the several different GX theories, there are concepts that consistently occur as important constituents of GX. One such concept is challenge. Indeed, without a challenge that is perceived as worthy, games are simply not enjoyed as much [16]. However, challenge, as used in this context, is that reported by the gamer on reflection of their experience. How this relates to the actual challenge presented by games has only begun to be examined [22] and even then the definition of challenge is purely in terms of players' perception of difficulty. It is possible that respondents to a questionnaire incorporating challenge elements are merely asserting the challenge experienced as some sort of halo response to a generally positive (or negative) GX. And even if this were not the case, none of the approaches are precise about what specifically the challenges are that lead to a particularly good (or bad) GX.

Designers of games necessarily consider how to make a game challenging to the gamer [23]. There are two main ways to achieve this: push the gamer's physical limits; or push the gamer's cognitive limits [14, 19]. The gamer's physicality limits the speed with which interactions with the game can be conducted, and the accuracy with which actions can be performed. The gamer's cognitive abilities have a limiting effect on speed and accuracy of the problem solving activities required by the game.

Immersion is believed to specifically arise from the level of challenge experienced by gamers. As such, challenge, as perceived by the player, is a component that goes into the make-up of immersion [15] and can be isolated as its own factor of immersion. However, as with other aspects of GX, what is undefined is what sort of challenge is required in order to increase the immersive experience. In this paper we explore how both the cognitive and physical abilities of a gamer can be challenged, and investigate which has the greater effect on the level of immersion experienced. In our first experiment we manipulate the number of interactions required to make progress in the game and thus the speed with which the gamer must interact with the game. Our second and third experiments manipulate the level of time pressure under which the participants play the game. In contrast to experiment one which simply manipulates the number of actions the participants have to make, the addition of time pressure in experiments two and three means that participants have to also think faster. We hypothesise that it is the cognitive challenge that will have the greater effect on immersion.

Expertise

Csikszentmihalyi [8] suggests that flow is achieved as a result of an appropriate balance between the perceived level of challenge and the person's skills. Given the relationship between immersion and flow, we are interested in whether immersion is also sensitive to the balance between skill and challenge.

Therefore, in addition to manipulating time pressure and number of interactions, we also explore the role of expertise in immersion. We hypothesize that, if the gamer finds the game too challenging, where this level of challenge is too high in proportion to their skills, they are more likely to experience anxiety and may 'give-up' playing the game. However if the inverse is true, whereby the gamers skills are too high and the game is not challenging enough, then the gamer will be in a state of boredom. When a gamer is in either of these states (anxiety or boredom), they are less likely to be immersed in the game. In two of the studies reported here, we therefore also explore the role of expertise in combination with challenge to see how the two together influence immersion.

Overview of the experiments

To experimentally investigate the role of challenge on immersion, it is necessary to manipulate games within the strictures of experimental control to provide the different sorts of challenge. The ideal is two versions of the same game that differ only in the physical or cognitive challenge that they present.

For the first experiment, manipulating physical challenge could have been done by simply requiring more actions to achieve the same effects, eg two mouse clicks to select an object rather than one. However, unless this is plausibly motivated by the game, players will perceive such naïve manipulations as unnecessarily awkward and immersion could be reduced simply because of frustration at poor game design. The aim for the game in the first experiment was therefore to provide essentially the same playing experience with the same level of difficulty but that simply required more actions to keep playing. As will be seen this manipulation does seem to have been successfully achieved.

For the second and third experiments that intend to manipulate the cognitive challenge, it is not enough to provide more difficult puzzles or tasks for players as such increases in cognitive challenge may leave the players unable to make any progress in the game and decreases could leave the game too easy to be interesting. Instead, we chose to manipulate the time pressure on the player. This would mean that players would be required to achieve the same cognitive tasks but at different rates. This contrasts with the physical challenge of the first experiment in that if players failed to meet the cognitive challenge in the second two experiments then the game would end very quickly whereas in the first experiment, failure to keep up with the physical challenge would not mean immediate game over and the players would have the opportunity to subsequently act quickly to make up for lost time.

Of course it must be acknowledged that in manipulating challenge whilst maintaining comparability between experimental conditions, it is difficult to completely disentangle physical and cognitive challenge. Nonetheless, in the first experiment the emphasis of the manipulation is on physical challenge whereas on the latter two it is on cognitive challenge.

It should also be noted that the first experiment was devised in parallel with the second and third experiments due to the circumstances of the researchers involved. Thus, the design of the second and third experiments did not directly build on that of the first even though they did conceptually. Naturally, all three experiments drew on the same background research.

EXPERIMENT 1: INVESTIGATING THE EFFECT OF SPEED OF INTERACTION ON IMMERSION

The first experiment investigated how altering the level of physical effort required from players influenced their sense of immersion within the developed game. This was altered by changing the relative game 'intensity', represented by the level of activity required to play the game.

A number of possible effects could have been observed. Compelling players to be more active may have caused them to focus more intently upon the game, increasing their attention and potentially immersion. Conversely, 'frantic' gameplay could overstress players, preventing them from appreciating other aspects of the game, decreasing the cohesiveness of their experience and thus reducing their overall immersion level.

In addition, we explore the effect of expertise by categorizing players according to their prior experience with the game genre, in this case, Tower Defence Games.

Method

Design

This was a between subjects design with two conditions. The independent variable was the relative amount of effort required by the game: low effort (LEff) and high effort (HEff). The dependent variable was the participant's sense of immersion within the game, as calculated through participant responses to the Immersive Experience Questionnaire (IEQ) [15].

Participants

41 students from the University of York took part in the study; however, one participant's results were removed due to inattention and lack of effort during the study. As a result, there were 40 participants for analysis (29 male). Ages ranged from 18 to 24 (mean = 20.9).

Game experience and attitudes varied between the participants, ranging from those who played games less than once per month, to game programmers and even the chair of the York University LAN (Local Area Network) gaming society.

Materials

The game used in this study was a Tower Defense (TD) game purpose built for research in GX in general and this study in particular. The basic idea is common to all TD games where enemies (creeps) move from one side of the screen to the other along a set path. The player must place towers along the path to damage, hinder and ideally destroy all creeps before they reach the other side. Creeps have various characteristics such as speed, tolerance to damage and so on. Towers also vary in how they affect creeps, how fast they fire and they can further be upgraded to improve their characteristics. Towers and upgrades cost money which is generated by killing creeps. Players start with 20 lives and one is lost each time a creep manages to cross the screen. Play stops if all lives are lost or after eight minutes when the creeps stop attacking.. Because the pattern of creeps is the same each time, a simple measure of performance is how long a player survives before 20 creeps make it across the screen.

The advantage of using a specially developed game is that we could exercise experimental control over the gameplay to achieve a manipulation of physical challenge whilst maintaining the level of cognitive challenge. To do this, we used the notion of the game being "balanced." That is, the game is balanced if the amount of damage per second that players can buy matches the amount of damage that is needed to kill creeps at the required rate. To achieve this, the towers and creeps were specified so that despite the apparent range of differences between firepower of towers and the strength and speed of the creeps, provided the player placed the towers reasonably well and continually spent all income earned from killing creeps there would always be enough firepower to destroy the creeps. The only cognitive challenge would be where to place the towers and to a large extent that would not matter too much.

The initial game setup was determined through testing to create an appropriate rate of play. This was then reduced by approximately one third to create the low effort condition (LEff), and increased by approximately one third to form the high effort condition (HEff). In the LEff condition, players earned small amounts of money sufficient to buy powerful towers whereas in the HEff condition, players earned more money to by more but weaker towers than in the LEff condition. In both conditions, the game was balanced.

It may be argued that the HEff condition required more cognitive effort because of the task of having to frequently buy or upgrade towers, that is, the apparent time pressure. However, this is not the case because to a large extent it did not matter how players spent their money, simply that they spent it. Also, there was no specific time pressure to act at any particular moment. Falling behind in spending could be compensated for by a rapid burst of spending at a later point.

The game was loaded in a resolution of 1440x900 (full-screen) on a 15.5" laptop with a 2.2GHz Intel Core2 Duo processor, 4GB DDR2 RAM and an NVidia GeForce 8600M GT graphics card. A basic, ambidextrous mouse was provided for input, and on-ear headphones were supplied to allow sound effects and music to be heard.

Immersion was measured using the IEQ [15] which consists of 31 questions. An additional 3 questions were added to the questionnaire to explore to game playing habits and how players rated their own ability. These questions were used in order to estimate player expertise.

Procedure

The game included an in-game tutorial, page one of which was loaded prior to participant arrival. Players were required to read, or at least navigate through the pages of, this tutorial before access to the actual game was provided.

Participants then played the game for either eight minutes or until the player had run out of lives. Upon reaching a game-over condition, players were asked to immediately inform the investigator that they had finished playing. Participants were then provided with the IEQ and again left alone to answer this privately. When they had completed the IEQ, participants were asked if they had any further questions or comments, before being orally debriefed.

Participants received a £5 Amazon voucher for taking part.

Results

The purposes of the different experimental conditions were to impose significantly different levels of physical effort whilst keeping difficulty constant. Table 1 shows the number of actions and performance (seconds survived) of participants in each condition. There was a significant difference between the number of actions performed in each condition ($F(1, 38) = 13.17$, $p < 0.001$). Participants in the HEff condition performed, on average, over 60% more actions than those in the LEff condition. There was no difference in performance between the conditions ($F(1, 38) = 0.032$ $p = 0.86$). These results demonstrate the success of the setup of each experimental condition in achieving the intended effects.

	Low Effort	**High Effort**
Actions	48.7 (19.8)	78.2 (30.5)
Performance (s)	391.1 (99.5)	397.1 (112.2)
Immersion	112.2 (15.6)	113.5 (12.7)

Table 1: Mean number of Actions, Performance and Immersion (and standard deviations) in the experimental conditions

Table 1 also shows the immersion levels of participants in each condition. The results obtained demonstrate no significant effect of physical effort upon immersion ($F(1, 38) = 0.09$, $p = 0.766$).

Expertise

Data analysis using Pearson's r (two-tailed) indicated a significant positive correlation between expertise and performance ($r(38) = 0.62$, $p < .001$), confirming the adopted measure of expertise was appropriate for this study. Participants were divided into three expertise groups using a three-way split. The lower group (N = 13) was termed 'insufficient expertise' (IX), the middle group (N = 16) 'low expertise' (LX) and the upper group (N = 11) 'high expertise '(HX).

Table 2 shows the performance of the different expertise groups in each condition. Players in the IX group performed significantly differently ($F(2, 37) = 12.78$, $p < 0.001$). Additionally, IX players performed a significantly different number of 'upgrade' actions relative to the number of total actions performed ($F(2, 37) = 13.19$, $p < .001$), with many performing no upgrades at all (38% against 0% in the other two groups) (see Table 3). As such, players from this group were missing out on important gameplay elements, indicating that they had not fully read or understood the provided tutorial and the basic mechanics of the game. For these reasons, it was decided to exclude IX players' results from further analysis as these were not representative of the gaming experience of typical game players.

	Level of Expertise		
	Insufficient (IX)	**Low (LX)**	**High (HX)**
Overall	299.6 (118.9)	431.3 (60.7)	451.7 (52.8)
LEff	328.1 (123.7)	415.9 (70.8)	439.6 (64.5)
HEff	266.3 (114.5)	446.6 (48.4)	461.8 (44.5)

Table 2. Mean (standard deviation) of performance in the different expertise groups

	Level of Expertise		
	Insufficient (IX)	**Low (LX)**	**High (HX)**
Actions	51.1 (32.1)	88.9 (22.2)	94.8 (16.5)
Upgrades	23.4 (28.1)	57.3 (19.4)	65.3 (18.9)

Table 3: Mean (standard deviation) of the number of actions and upgrades in the different expertise groups

The immersion levels of players in the LX and HX groups are shown in Figure 1. Immersion levels are consistent across both groups in the LEff condition (LX = 115.5 (s.d.=14.5), HX = 115.0 (s.d.=12.7)). However in the HEff condition, immersion levels increase for the LX group (123.0 (s.d.=8.7)), but decrease for the HX group (104.2 (s.d.=9.6)). These results show a statistically significant main effect of expertise ($F(1, 23) = 4.48$, $p = .045$) and an interaction between expertise and physical challenge at a level approaching significance ($F(1, 23) = 4.03$, $p = .057$).

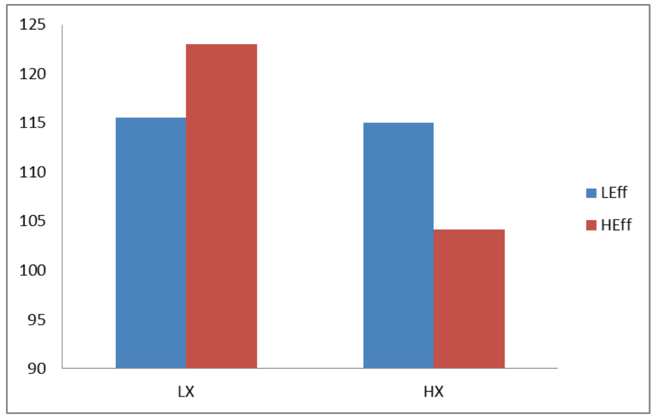

Figure 1: Mean immersion scores in the low and high expertise groups

	Low (LX)	**High (HX)**
LEff	14.1 (3.44)	13.2 (2.16)
HEff	14.9 (2.75)	11.0 (2.83)

Table 4. Mean (standard deviation) of challenge in the different expertise groups

Analysing just the challenge factor of immersion, the means and sd for challenge are given in Table 4. The pattern of results is similar to that for overall immersion but there is a main effect of expertise ($F(1,23) = 4.42$, $p=0.047$) but no interaction effect between expertise and physical challenge ($F(1,23)=1.67$, $p=0.209$).

Discussion

The experimental setup had the desired effects. The number of actions performed increased dramatically and, most importantly, significantly in the high effort condition, indicating an increased level of physical effort being placed upon the participants in this condition. The consistent performance of players' shows that this difference is not as a result of the amount of time spent playing in either condition, demonstrating conditions were equal in difficulty and supporting the requirement that the increase in the level of physical effort required does not impact performance.

Despite this success in manipulating the number of actions required to play the game, but not the game difficulty, the observed levels of immersion were consistent between conditions, demonstrating that physical effort alone does not impact significantly on the level of immersion experienced.

Where players are actually able to fully engage in the game, that is discounting 'insufficient expertise' players, it is clear that those with low expertise view the two experimental conditions equally. However, those with more experience are less immersed in the HEff condition than the LEff despite performing at a similar level. The Low expertise players have a fine time in the Leff condition, and a slightly better time when pushed to go faster. However, for those with more experience and skill (HX), the requirement to play faster is frustrating as the speed with which they must complete actions in the game negatively impacts on their abilities to apply their strategies. Indeed, some of these players commented in debriefing that they did not have time to work out which towers were best to deploy.

The same picture is seen when considering only the challenge factor with high expertise group in the HEff condition rating the challenge lowest. This may seem contradictory: the HEff condition should produce higher challenge. It does however support the view that the questionnaire, measures the player's perception of challenge in the context of the overall gaming experience, and that, when evaluating the GX, the gamer values cognitive challenge rather than physical challenge (at least in this game genre).The low scores on the challenge factor suggest that the High expertise players found the HEff condition not to be challenging perhaps because it was an unfair or unreasonable challenge. When a player is an expert, the challenge of this game is cognitive (identifying which towers are best from a strategic perspective) not physical.

EXPERIMENT 2: INVESTIGATING THE EFFECT OF TIME PRESSURE ON IMMERSION

For experiment 2, the aim was to investigate the effect of time pressure on gaming. In contrast to the manipulation used in experiment one, we hypothesized that the addition of time pressure would manipulate the level of cognitive challenge within the game. This is because the player is not only required to complete more actions per time unit in order to play the game, but that they also have to think more quickly in to stay in the game and maintain immersion. It was therefore predicted that participants who play the experimental game under time pressure will have significantly higher overall immersion scores and 'challenge' scores than those who play the game under no time pressure. As the focus in this study was on the explicit link between challenge and immersion, expertise was not considered in relation to challenge. Instead of taking statistical account of it, it was experimentally controlled by balancing the levels of expertise between the experimental conditions.

METHOD

Design

The study was a between subjects design. The independent variable was the level of challenge in the experimental game, i.e. whether the game had time pressure or no time pressure. The dependent variable was the participant's questionnaire scores on the IEQ.

Participants

There were 22 participants (12 male) from University College London with an age range of 18-22 years old and an average age of 20.1 years.

Materials

The PC game Bejeweled [2], was used in both conditions. The game can be played in either "simple" or "timed" mode. The aim of the game is to swap adjacent gems to align sets of 3 or more gems of the same colour. Scores are dependent on gems being aligned and sets being created.

The timed mode used for the time pressure condition has a Gem Meter which displays the amount of time the player has left to play shown by the length of the green line on the meter. To prevent time running out, the player must keep finding as many sets of the same gem colour as possible. The more sets found, the more power the Gem meter obtains and this in turn enables bonuses as well. If, however, the player has long intervals between finding sets, the length of the gem meter would gradually decrease so long as no sets are being made. As the game progressed, the participant would find it more difficult to make sets of gems and the Gem meter would decrease more quickly. If the participant failed to make any sets, then their time would eventually run out and the game would be over.

The simple game, used for the no time pressure condition, also has a Gem Meter, however the meter increases in power if sets are made but does *not* decrease in power if the player is unable to make sets of gems. Participants are therefore not playing against time. The game is completed when there were "no more moves left" i.e., when sets could no longer be made.

The IEQ [15] was used to measure the levels of immersion experienced.

Procedure

Participants first filled out a questionnaire that asked questions about their general gaming experience and particularly their experience with the experimental game Bejeweled. This enabled us to further divide the participants in each condition so that half of the participants would have played the experimental game and half would not have. The participants were then given a set of standardized instructions, which were specific to the experimental condition.

Participants were required to play the game Bejeweled. Those in the time pressure condition played the 'timed' version of the game and those in the no time pressure condition played the 'simple' version of the game.

The participants were asked to play the game for approximately 15 minutes in each condition unless the game ended before this. Participants then completed the IEQ.

Results

There was no significant difference in the scores achieved by participants in the two groups (time pressure mean score = 5603 (s.d.=6212), no time pressure mean score = 3501 (s.d. = 3187) $t(20)=0.998$, $p>0.05$).

The descriptive statistics presented in Table 5 clearly show that those in the time pressure condition have a much higher immersion score than those in the no time pressure condition. This was confirmed by paired samples t-tests: overall immersion scores $t(10)=6.36$, $p<0.05$); challenge $t(10)=9.59$, $p<0.05$.

Questionnaire Scores	Time Pressure	No Time Pressure
Total immersion*	113.5 (11)	81.3 (8.60)
Challenge*	15.6 (1.12)	8.82 (1.78)

Table 5. Mean (standard deviation) of the overall immersion score and the challenge factor score.

Discussion

There was no difference in the scores in the two conditions, thus any differences found in levels of immersion cannot be attributed to one group being better than another.

The results and analysis for experiment 2 clearly show a significant difference between the conditions on both the overall immersion questionnaire scores and the scores for the challenge factor. This confirms our hypothesis that it is the addition of cognitive challenge that impacts on the GX.

In experiment 1 we argued that those with insufficient experience were not representative of the gaming experience of typical game players. The participants in experiment 2 fell in to two broad categories, those who had played Bejewelled before, and those who were complete novices. We therefore have not analysed differences between participants with different levels of experience, but simply ensured they were equally distributed between the experimental groups. In study 3 we therefore explore the role of expertise in the context of time pressure.

EXPERIMENT 3: INVESTIGATING THE INTERACTION BETWEEN TIME PRESSURE AND EXPERTISE ON IMMERSION

The aim of this study was to test the hypothesis that expertise interacts with the level of cognitive challenge imposed by the game and thereby mediates the level of immersion experienced. We manipulated this by selecting participants who were either expert or novice at a particular game (Tetris), and tested how immersed they were whilst playing the game, that was either high in challenge or low in challenge. In this study, the challenge of the game was determined by the level at which the participant was playing the game, so the higher the level, the more challenging the game. The difference in levels is in the speed at which the game progresses, thus making this manipulation analogous to that in study 2.

It was thus hypothesised that 'experts' will be more immersed when playing the experimental game at a higher level, but will be in a state of boredom when playing the game at a lower level. It was also predicted that 'novices' will be more immersed when playing the experimental game at a lower level, but will be in a state of anxiety when playing the game at a higher level.

Method

Design

The study was a mixed design, which was carried out in a laboratory setting. The within subjects independent variable was the level of the experimental game (Tetris) that was being played (which was either low- level 1 or high- level 6). This would indicate the level of challenge within the game. The between subjects variable was the level of expertise (expert or novice). The dependent variable was the participant's questionnaire scores on the immersion questionnaire (IEQ).

Participants

There were 20 participants (10 male) from University College London. Their age range was 18-26 years old and the average age was 19.3.

Materials

The PC puzzle game Tetris [28] was used in both conditions. The aim of Tetris is to make as many lines as possible by stacking up the pieces that fall one by one from above. The pieces are stacked accordingly, by either rotating their positions (to ensure a correct fit) or by moving the piece towards the left or the right of the screen (depending on where the player wants to place the piece). In order for the player to make the lines, the pieces must be packed tightly so that there are no 'gaps'. As they progress through each level, the speed at which the pieces fall from above will increase, making it more difficult and more challenging for the gamer to decide how and where to place the piece in the stack.

Additionally, participants were given a questionnaire that asked about their gaming experience. This was a similar questionnaire to that used in study 2, which provided information as to whether the participant plays and how often they play computer games, the kind(s) of computer games they enjoy playing and kind(s) they find challenging. It also asked the participants whether they had played the experimental game Tetris before. If they answered 'yes' to this question, they were asked how often they play the game and were also asked to give an honest rating of how 'good' they think they are at it on a scale of 1-10 (where 1= novice and 10= expert). This gave us an indication of their abilities and skills regarding the game and therefore enabled us to assign the participants to each condition so there would be 10 participants who rated themselves closer to being 'expert' at playing Tetris and 10 participants who rated themselves closer to being 'novice'.

Immersion was measured using the IEQ as in the previous studies. An additional questionnaire asked participants to rate their levels of boredom and anxiety.

Procedure

Participants first completed the gaming questionnaire. They then played each condition (at level 1 or at level 6) for approximately 15 minutes. Order of play level was counterbalanced between participants.

When in the level 1 (the low challenge) condition, they were allowed to complete levels 1 and 2 but once they had reached the third level, they were asked to restart the game. They were asked to do this for the entire duration of the gaming portion of the study (restarting the game whenever they reached level 3).

The participants in the level 6 (the high challenge) condition were allowed to play the game until the game ended, i.e. when the stack had reached the top and no more lines could be made. Once their game was over, they were asked to restart the game at level 6. They too were asked to do this for the entire duration of the gaming portion (restarting the game at level 6 whenever the game ended).

Once participants had completed each condition of the game, they were asked to complete the IEQ. Participants were also asked to rate how bored or anxious they were (on a scale of 1 to 10) after playing in each condition.

Results

When playing Tetris at a high level experts were more immersed (mean = 110.4, sd = 12.68) than when playing at a low level (mean = 81.9, sd = 12.34). The novices on the other hand, were more immersed when playing Tetris at a low level (mean= 106.2, sd = 9.93) than when playing at a high level (mean = 83.2, sd = 17.04).

There was a highly significant interaction, ($F(1,18)=53.37$, $p<0.001$). As the level of challenge increases, experts become more immersed in the game and novices become less immersed in the game. The repeated measures ANOVA however, also showed that there was not a significant main effect of level of expertise ($F(1,18)=0.93$, $p>0.05$) or of level of challenge ($F(1,18)=0.61$, $p>0.05$), suggesting that the level of expertise in the game (expert or novice) and the level of challenge (low or high) do not vary significantly when expertise is unaccounted for.

	Low	High
Expert	7.8 (2.44)	13.9 (1.28)
Novice	14.7 (2.41)	14 (2.21)

Table 6: Descriptive statistics of the challenge factor scores of the low level and high level conditions.

Table 6 clearly shows that the experts mean challenge score for the high level condition was almost double that of the low level condition, suggesting that the experts found the high level condition to be far more challenging than the low level condition.

The novices however, had a slightly higher mean challenge score in the low level condition than in the high level condition. The mean challenge scores for both conditions are extremely high (given that the maximum possible challenge score is 20) with a difference of just 0.7 between the low level conditions.

Discussion

Whilst the effect of challenge shows no influence on immersion when there is no consideration of expertise, when expertise is factored in, there is a significant interaction showing that expertise and challenge interact in an important way in leading to immersion. This result mirrors the well-known effect in flow whereby it is only through balancing skill and challenge that flow is achieved. Immersion however is a much more general experience than flow and moreover is graded so that people can assess themselves to be more or less immersed. Nonetheless, the same pattern is seen where those who are less experienced are more immersed at lower levels of challenge and conversely those with more experience are more immersed at higher levels of challenge.

GENERAL DISCUSSION AND CONCLUSIONS

The final experiment might give the impression that immersion behaves like flow with a balance of challenge and expertise leading to better gaming experiences but this needs closer scrutiny. The classic view of flow is that achieving the balance of challenge and expertise leads to the optimal experience, flow, and regardless of how expert you are, the balance gives the same top end flow experience. Here though all players achieve some degree of immersion in all of the games. Immersion is higher though when the balance of challenge and expertise is better. This then supports the idea that as players become more immersed they move towards total immersion or flow [25].

Additionally, this picture of increasing immersion is in contrast to the situation in the first experiment. The game of Tetris offers little subtlety: blocks fall and the only challenge is posed by how fast they fall. Thus experts at Tetris see the game, know the game and their expertise is challenged at higher levels. By contrast, tower defense games offer a variety of challenges where success requires strategic thinking and games vary substantially in which strategies are successful. High expertise players know this and require time to assess the relative strengths and weaknesses of the towers in order to deploy them optimally. However, the tower defense game in the first study did not require this level of strategic thinking.

These results suggest that whilst the balance of skill and challenge is important, it is also important that players perceive themselves as having the suitable skills. With Tetris, this was possible and so the higher challenge was more immersive. With the tower defense game of the first study, this was not so straightforward and the higher skill players recognized their inability to assess the challenge and therefore their skills in relation to the challenge. These players were less immersed being less able to deploy their higher skills against the higher challenge. Lower expertise players however, unaware of the challenges they faced were much more able to just get on and play the game and so were able to be equally immersed regardless of the physical challenge they faced.

High expertise is not only a matter of high skill levels but also the ability to use expertise to reflect on the matters at hand [24]. The effects seen here present something of a converse of the Dunning-Kruger effect [17]. For the Dunning-Kruger effect, those with low expertise were unable to properly understand feedback and therefore failed to re-evaluate themselves appropriately in light of feedback. Those with high expertise better understand feedback and so re-evaluate themselves more accurately in response to feedback. Here by contrast, those with low expertise just got on with playing and enjoying the game whilst those with those with high expertise mis-perceived the challenge presented due to the shallow nature of the game and therefore enjoyed themselves less.

The first and second study together also show that whilst it is important that players have activities to engage them, it is not simply a matter of having stuff to do. Immersion does not arise from physical challenge but from the cognitive challenge of solving problems quickly not simply acting quickly.

Our findings are novel in the context of GX. Whilst the many approaches to GX acknowledge the role of challenge in leading to positive gaming experiences, there has been little attempt to qualify what form these challenges must make and how they relate to players expertise. The studies here present a first attempt to provide construct validity to the notions of GX through explicitly manipulating the concepts believed to underpin GX from an objective perspective of what constitutes challenge.

Of course, this work only begins to scratch the surface of the types of challenge that games offer and the myriad of GX that can therefore arise. Further work should not only explore challenges arising at the basic level such as time pressure but also things that lead to higher scores or collecting more bonus points. In addition, this research suggests that challenges at higher levels such as the level of strategy and possibly even narrative are important for immersion. There is always concern for the limitation of lab studies such as these and what challenges players expect over a longer term engagement with games are currently unspecified.

Throughout all of this, it is also clear that challenge cannot be considered in isolation but must always be analysed in relation to the expertise of players. Even this is not simple in that challenge is distinct from the perceived challenge as indicated in the first study so that when attempting to achieve immersion through a balance of challenge and skill, players must first be able to perceive that such a balance is possible.

ACKNOWLEDGMENTS

We thank the EPSRC for supporting the first experiment on grant TS/G002843/1. And we would like to thank all our participants for giving their time to play games for us, and the reviewers for their insightful comments that improved this paper.

REFERENCES

1. Brockmyer, J., Fox, C., Curtiss, K., McBroom, E, Burkhart, K., Pidruzny, J. (2010) The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *J. of Exp. Social Psychology*, 45(4), 624-634.
2. Bejeweled: http://www.1001onlinegames.com/games/puzzle/bejeweled.html
3. Brown, E. & Cairns, P.A,, (2004). A grounded investigation of game immersion. *CHI 04*, 31-32.
4. Calvillo-Gamez, E., Cairns, P., Cox, A.L. (2010) Assessing the core elements of the gaming experience. In Bernhaupt, R. (ed) *Evaluating user experience in games*, Springer, 47-71
5. Chatfield, T. (2009) Videogames now outperform Hollywood movies. *The Observer, Sunday 27th Sept 2009*
6. Chen, J. (2007) Flow in games (and everything else). *Comm. ACM,* 50(4), 31-34.
7. Cowley, B., Charles, D., Black, M., Hickey, R. (2008) Toward an understanding of flow in video games. *ACM Computers in Entertainment,* 6(2), 20:1-20:27
8. Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience.* New York: Harper and Row.
9. de Kort, Y. A. W., IJsselsteijn, W. A., Poels, K. (2007). Digital Games as Social Presence Technology: Development of the Social Presence in Gaming Questionnaire (SPGQ). *Presence 2007*.

10. El-Nasr, M.S., Aghabeigi, B., Milam, D., Erfani, M., Lameman, B., Maygoli, H., & Mah, S. (2010) Understanding and evaluating cooperative games. *CHI '10. ACM, New York, NY, USA*, 253-262
11. Epps, A. (2007) *Immersion in video-games: How is immersion related to attention?* Unpublished MSc Thesis, UCL
12. Ermi, L., Mayra, F. (2005) Fundamental Components of the Gameplay Experience: Analysing Immersion. *Proc. of DIGRA 2005*
13. Gagne, A., El-Nasr, M.S., & Shaw, C. A Deeper Look at the use of Telemetry for Analysis of Player Behavior in RTS Games. *International Conference on Entertainment Computing (ICEC 2011), 2011*
14. Hsu, S.H., Wen, M.H., Wu, M.C. (2007) Exploring design features for enhancing players challenge in strategy games. *Cyberpsychology & Behaviour* 10 (3), 393-397
15. Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, 66(9), 641-661.
16. Juul, J. (2009) Fear of failing? The many meanings of difficulty in video games. In Wolf, M and Perron, B. (eds) *The Video Game Theory Reader 2.* New York: Routledge, 237-252.
17. Kruger, J., Dunning, D, (1999) Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, Vol 77(6), 1121-1134
18. Livingston, I., Nacke, L., Mandryk, R. (2011) Influencing Experience: The Effects of Reading Game Reviews on Player Experience. *International Conference on Entertainment Computing (ICEC 2011), 2011*
19. Orvis, K.A., Horn, D.B., Belanich, J., (2008) The roles of task difficulty and prior videogame experience on performance and motivation in instructional videogames. *Computers in Human Behaviour* 24 (5), 2415-2433
20. Poels, K, deKort, Y, IJsselsteijn, W. (2007) "It is always a lot of fun!" - exploring dimensions of digital game experience using focus group methodology. *FuturePlay 2007,* 83-89
21. Qin, H., Rau, P.L.P., Salvendy, G., (2009) Measuring player immersion in the computer narrative. *International Journal of Human-Computer Interaction* 25 (2), 107-133
22. Qin, H., Rau, P.L.P., Salvendy, G., (2010) Effects of different scenarios of game difficulty on player immersion. *Interacting with Computers* 22, 230-239
23. Schell, J. (2008) *The Art of Game Design.* Burlington MA: Morgan Kaufmann.
24. Schon, D (1983) *The Reflective Practitioner.* London: Basic Books
25. Seah, M., Cairns, P. (2008) From Immersion to addiction in videogames. In England, D. and Beale, R. *Proc. of HCI 2008*, vol 1 BCS, 55-63
26. Skalski, P., Tamborini, R, Shelton, A., Buncher, M, Lindmark, P. (2011) Mapping the road to fun: Natural video game controllers, presence, and game enjoyment. *New Media & Society,*
27. Sweetser and Wyeth. 2005. GameFlow: a model for evaluating player enjoyment in games. *Comput. Entertain.* 3, 3 (July 2005), 3-3.
28. Tetris: http://www.capehostpro.com/play-tetris.htm
29. Wallop, H. (2009) Eight out of ten homes own a next-gen console. *The Daily Telegraph, 15th Jan 2009*
30. Wijnants, M., Agten, S, Quax, P., Lamotte, W. (2009) Investigating the relationship between QoS and QoE in a mixed desktop/handheld gaming setting. *CoNEXT Student Workshop '09,* 29-30