Conference contribution :
Denisova, A. & Cairns, P. (2015). *The Placebo Effect in Digital Games.*(pp. 23-33). London, United Kingdom: Symposium on Computer-Human Interaction in Play (CHI Play).
http://dx.doi.org/10.1145/2793107.2793109

The Placebo Effect in Digital Games: Phantom Perception of Adaptive Artificial Intelligence

Alena Denisova
Department of Computer Science
University of York
York, YO10 5GH, UK
ad595@york.ac.uk

Paul Cairns
Department of Computer Science
University of York
York, YO10 5GH, UK
paul.cairns@york.ac.uk

ABSTRACT

Play-testing of digital games is a crucial part of any game development process, used to gather feedback about the game and correct any existing and potential flaws with the design. However, due to the nature of human subject testing, the feedback being collected in such experiments is prone to biases. Players' expectations play a great role in dictating their gaming experience, which means the information players receive before trying a new game, as well as the knowledge they already possess, may affect their perception and experience of the game. Two studies were conducted in order to evaluate how priming players to expect improved technology can positively influence their experience. The results supported the hypothesis that even basic instructions can change players' perception of the game, and lead to a higher level of perceived immersion when knowing that the game contains an improved feature, the adaptive artificial intelligence (AI), while it is not present in the game.

Author Keywords

Player experience; Immersion; Player knowledge; Biasing effects; Adaptive artificial intelligence.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

When it comes to understanding the player experience (PX) of digital games, of course, the game itself is crucial for establishing the experience: playing *Mario Kart* is very different from playing *Silent Hill*. Research backs this up with in-game factors, such as the presence of music [24] and the level of challenge [8], having a measurable effect on PX. It is also becoming well understood, though, that factors external to the game can also have an important influence. These include the size of the screen [28], surrounding lighting [22], and even the personality of a player [1].

CHI PLAY 2015, October 03 - 07, 2015, London, United Kingdom.

ACM 978-1-4503-3466-2/15/10 ...$15.00
http://dx.doi.org/10.1145/2793107.2793109

Given such a wide range of influences on PX, it can be hard to know to what extent the experiences of a player are due to the game or the player's expectations about the game. Game reviews [20] and ratings [14] are capable of changing players' attitude towards a game before they have even played it. However, reviews and ratings are based on subjective opinions of other gamers and gaming experts, which mean a great deal in shaping opinions of new players before they get to try the game themselves. These cognitive biases, namely conformity to majority [26] and conformity to authority [27], often arise from reading such information. On the other hand, players often expect the upcoming game to have new features and novel technology, which can be important in making a good first impression for many players. These expectations of a game, which can be affected by external information, are therefore increasingly important in terms of shaping gaming experience, and matching them in the game is believed to lead to a heightened sense of immersion [21]. The question is therefore to what extent expectations alone can influence the player experience?

This then becomes a further problem in the context of play-testing, a mainstay of modern game development processes [25]. Play-testing is used to debug problems in the game, be it in the code or in the game design, but also to provide some sort of summative evaluation of the experience that players will get [13]. However, if players know they are testing some new version of a game, that alone may be enough to influence their experience as they may have expectations of a novel digital game technology, improved gameplay and so on. This can be independent of any explicit prompt from the game developers keen to seek honest and unbiased feedback.

In this paper, we therefore consider the impact of simple expectations on PX. To do this, we draw on the placebo literature where people's expectations of medical procedures has been consistently shown to have profound effect [12]. Of course, it is easy for people to report more fun and enjoyment in a game based on expectations, we therefore have chosen to focus in more on the actual engagement players have, as typefied by their experience of immersion [5]. Immersion is one of the most common terms used to describe a feeling of being highly involved in a digital game [15] not only in the gaming communities but also by the digital game researchers [5]. Amongst other PX theories and models, the broad nature of the term immersion allows for the freedom of definition, which makes it possible to account for the human aspects of the concept. We report on two studies that show how telling players about presence of a certain technology in the game, though not necessarily telling them about the benefits of such technology, is sufficient to increase their immersion in a game. This both reveals the complexity of reliably producing player experiences and also has implications for the use of experiential measures in play-testing.

PLAYER EXPECTATIONS AND EXPERIENCE

As digital games become an integral part of our everyday lives, the need for understanding the experience of their users is becoming more prominent, resulting in various theoretical concepts that are used to describe what players feel and think. Though the experience of digital games is subjective and differs from player to player, there are certain types of experiences that remain the same for millions of players across a wide rage of games: entertainment, challenge, engagement, etc. In the analytical research of this field various terms have been established to try to account for these experiences. Some of the most widely used theories include flow [9], presence [30], engagement [4], player experience of needs satisfaction (PENS) [23], and immersion [15], each of which can be quantified using concept-specific questionnaires. Broadly, these concepts are used to describe PX, though they tend to differ in the context of use based on the breadth and depth of the concept and its fine details.

Immersion is one of the most frequently used terms by gaming communities and researchers to describe what players feel when being engaged in a game play [5]. It is an experience of deep involvement with a digital game [15], during which players forget about their everyday concerns, lose track of time and become less aware of their real world surroundings. To measure this experience, Jennett et al. [15] developed a questionnaire, which allows to estimate the overall level of immersion in the game. The questionnaire was extensively validated using a large scale survey and experiments, providing good support for the construct validity of the IEQ. The factor analysis of the IEQ, conducted as part of the validation, also suggested that immersion could be considered as having five constituent factors contributing to this experience – player's cognitive and emotional involvement, their real world dissociation, and challenge and control in the game, as they perceive it.

Due to the broad nature of its definition, immersion is often mistakenly assumed to be synonymous to the concept of flow, or being incorporated in this experience. However, immersion is a graded experience [5] unlike the optimal state of flow, in which players become fully absorbed in their activity so that nothing else seems to matter. In their chapter, Cairns et al. [6] provide a thorough comparison between these experiences, and several others, outlining the clear distinction between immersion and other gaming experiences.

PX is taken seriously by game developers, as the success of any game depends of what the player thinks. Nowadays, players' opinions are frequently shared online, where these positive or negative reviews help shaping the mind of a prospective player. Whether people would want to try a new game often depends not only on their personal preferences and previous experience, but also on players' expectation of a game, based on reviews they read and recommendations they receive from others. Players' immersion in the game is believed to be dependent on the appropriate match between the game world conventions and the expectations of the player [21].

The evidence that game reviews can influence players' perception and enjoyment of a digital game was found by Livingston et al. [20]. In their study the researchers compared the effect of positive and negative reviews on PX, and found that valence of review text affected game ratings. Similarly, Jenkins et al. [14] demonstrated that knowing about game ratings also tends to influence PX. Biases like these are common – players tend to be suggestible and therefore, any information they receive before playing a game for the first time can affect their unbiased opinion.

Due to natural biases, players' opinions about a game are likely to be affected by the play-testing instructions. The amount and quality of information provided by the experiment facilitator, the way it is presented to the participant, and the phrasing used to describe the game or particular features being tested, can bias the player. Moreover, people are susceptible to conformity, as in the case of reviews: if others think that a game is good, then an individual will adopt the same view [26]. Similarly, if an expert, for instance a play-tester, says a game is good, an individual tends to agree. This is called the authority heuristic [27]. Players also bring along their experiences of previous versions of a game, or games that are similar to them. In the former, players expect the new version of a game to be at least as good as the previous one, and even better, based on the previous experience. While experiencing a brand new game is different, as players are unable to compare it to some previous interactions with this or similar games. This could potentially cause a different perception of a game being play-tested. Players, who are trying out a new game, whilst being informed about new features, could concentrate on these features and through subconscious biases experience them as a substantial improvement. This is precisely the basis for the placebo effect in medicine.

THE PLACEBO EFFECT IN DIGITAL GAMES

Just like the placebo effect in medical and scientific trials depends on patients' expectations, experimental testing in a media domain can cause a similar reaction. The anticipation of the suggested reaction is said to lead to the generation of that reaction [11]. Expectations are powerful enough to motivate patients to improve their own medical conditions, or to motivate healthy people to increase their productivity and improve their performance. Hence, it is reasonable to suggest that the anticipation of a certain effect in games can lead a player to experience this effect. If an individual plays a digital game that is suggested to have a feature that may potentially improve their performance or the overall experience of the game, the player would expect it to happen. Believing in this will cause the player to subconsciously work towards achieving better results, or think that they enjoy the game more than if they were not aware of this feature.

Just as presentation is important for the patients to get better (e.g. medications' colours [2] and quantities [10]), the presentation is necessary for convincing the players that a particular feature in a game will improve the PX. The quality of explanation of the feature also has an effect on players' perception of it. Descriptions, which use technical language, are often seen as better quality due to their rather 'scientific' look. Because people do not often come across technical terms in everyday conversations, this may affect their evaluation of the explanations' logic [29]. People also tend to rate longer explanations as more similar to experts' explanations [18]. Moreover, non-experts are more easily convinced that the 'scientific' information they are reading is true and provides them with a good explanation of a topic, while people with more experience and expertise may be more sceptical about the validity of the provided description, and would not be allured by the technical presentation [29].

Based on the previous discussion, it is proposed that it is possible to alter the player's experience by suggesting that the game they are playing contains a feature can affect their performance. We chose adaptive AI as the feature for various reasons. First, it seemed to be a concept that players would recognise. AI is regularly referred to in games as how the game responds to players, typically with non-player characters, and it is known that companies are exploring these possibilities as a way to ensure playability by a wider-range of players [7]. By focusing on differences in individual learning and playing techniques in-game, instead of gathering data about players' requirements up front in the design process, it is possible to avoid issues typically encountered when using stereotypical data about players' sex and age [17].

Secondly, the adaptations that might be made are also understandable to players. Adaptive game technology could be used to provide appropriate challenge levels for each player. Algorithms are used to gather and learn information about the gaming patterns of a player in order to respond accordingly, e.g. adjusting an artificial opponent's strategy or varying the amounts of attainable items depending on player's progress in the game [16]. This can help players avoid getting stuck, modify gameplay to one's preference, or possibly identify if a player cheats or modifies the gameplay to their advantage [7]. When a player feels that the game is responsive to them as an individual, they may experience a heightened sense of immersion in the game world during game play. Moreover, one of the goals of adaptive AI design is to find a balance between the player's skill and the level of challenge of the game, which keeps the player in the state of flow [9].

In terms of managing players' expectations of the gaming experience, unlike improved graphics or sound, adaptive technologies are invisible to the player. As the game AI adapts to the player's individual approach, the player feels moderately challenged, and as a result feels increasingly immersed in the game. However, if the game does not have this feature, it is plausible that players could believe that it is present, and hence induce a realistic expectation of what the gaming experience should be like.

Thus, in the context of play-testing digital games, players could experience some kind of placebo effect, when explicitly told to expect adaptation in the game, which does not actually present there. It is hypothesised that the players who are given the game 'with adaptive AI' will feel more immersed in the game, than the players who are given the game 'without adaptive AI'.

STUDY I: WITHIN-SUBJECT DESIGN

The aim of the study was to determine the effect of information received from the instructor and players' knowledge prior playing a game on their PX. That is, whether suggesting that a digital game having adaptive behaviour in order to match the player skill set would affect their overall experience of the gaming session, particularly if they are comparing two sessions of the same game – 'with' and 'without' potential manipulation. In this study, which was a within-subject design, we looked at the behaviour and reactions of players, who compared two gaming sessions, in which one was supposedly with adaptive AI, and another had a 'standard AI'.

Experimental Method

Overall, 21 participants were recruited for the study – 3 female and 18 male students at the University of York. The average age of the participants was 23 (sd = 3.4), with the youngest player being 19 and the most mature one, 32. The majority of participants were regular gamers, spending more than one hour a day playing favourite digital games, several times a week. Those players listed strategies and role-playing games (RPGs) as the genres they prefer, although first person shooters (FPS) and sports games were also amongst the less frequently mentioned games. There were three people who stated that they do not often play digital games, but when they do, they normally spend over an hour in a single gaming session. These participants listed puzzle games, life simulations and racing games as their favourite genres.

Sixteen participants were familiar with the concept of adaptive AI, and 9 also had had previous experience of playing digital games that adapted to their behaviour.

Materials

The game chosen for this experiment was an action-adventure game with survival elements, *Don't Starve* [19], in which the player starts off with the main character, Wilson, placed in the middle of a procedurally-generated map with an empty inventory. The aim of the game is to collect objects in order to survive. The tasks include building new objects from the collected ones, and using them to protect the character from monsters that are randomly placed on the map, and natural events, such as weather and darkness. Random maps were used to avoid learning effects between the two conditions, and to avoid confounds due to idiosyncrasies of any particular map.

Unlike most commercially available digital games, 'Don't Starve' had the potential to be less known by frequent players, therefore reducing the chances of players comparing their experience of the game during the experiment to their previous experience. For the study, only participants with little to no experience of the game were able to take part.

The PX data was collected using the Immersive Experience Questionnaire (IEQ) [15]. An additional questionnaire was

Figure 1. Don't Starve *(2013) Klei Entertainment.*

used to measure the players' subjective views on the two gaming sessions – participants compared the level of enjoyment between the rounds 'with' and 'without' adaptive AI, and evaluated their performance in each round by answering eight Likert scale questions. The questionnaire was completed with an open-ended question about what participants believed the adaptive AI did in the game.

Procedure

The experiment started with a demographics questionnaire to gather the data about participants' gaming background. This was then followed by a short tutorial, during which participants played the game for about five minutes in order to familiarise themselves with the aim, the controls, and the storyline of the game. For the main part of the study, participants played the game two more times for 20 minutes during each session, counter-balanced for experimental condition.

All participants were given a brief explanation of what adaptive AI is before they played the game twice. The round 'with adaptive AI' was described to them as follows. Unlike in a standard game, the world was generated based on the participant's performance in the game – depending on how well they performed in the game, it would 'adapt' the game world to match the player's skills. The explanation given to the participants was deliberately vague in order to allow for the freedom of their imagination to explain what they experience.

A script was used to ensure that all players receive the same description of the feature being tested in the game, in order to avoid biasing participants. After each round of playing the game, they filled in the IEQ. Upon completing both sessions, participants filled in the final questionnaire collecting data about their experience of adaptive AI in the game. They were then fully debriefed.

Results

The main hypothesis was that there would be a difference in PX when playing a game 'with adaptive AI' and 'without adaptive AI'. Overall, participants felt significantly more immersed while playing the game 'with adaptive AI', in comparison to the session 'without adaptive AI' as determined by one-way ANOVA, ($F(1, 20) = 7.88, p = 0.011$), with a medium effect size – η^2_{partial} = 0.283. Figure 2 shows the difference between the total immersion in each session.

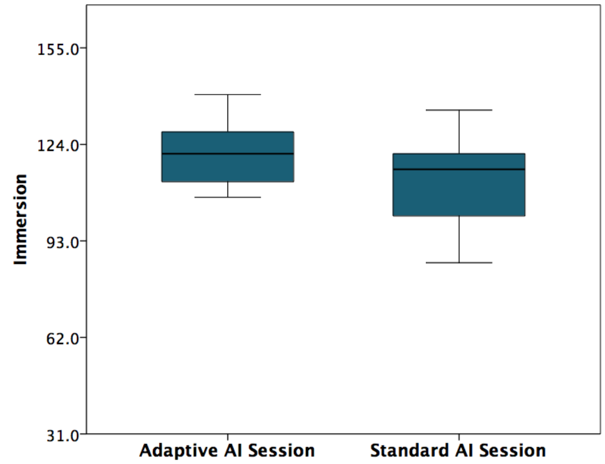

Figure 2. Total immersion level when playing 'with' and without adaptive AI.

According to Jennett et al. [15] immersion can be further broken down into five components – cognitive and emotional involvement, real world dissociation, control and challenge. Each of these elements were inspected in order to understand the obtained results.

The difference between the two sessions for each component is summarised in Table 1. Significant difference between the two sessions was found for two of these immersion components. During the 'adaptive AI' session participants felt that they were more cognitively involved with the game and were more cut off from the real world. There was no difference in terms of perceived challenge between the two gaming sessions, but there was marginal difference in terms of control participants experienced and their emotional involvement with the game during the 'adaptive AI' round.

Previous Experience

Whether participants were familiar with the concept of adaptive AI prior to the experiment appeared to affect their experience of the two sessions. People who had never come across this term before the experiment were more likely to believe in suggestions used in the experiment. The results from two-way ANOVA confirmed that the difference in players' immersion levels between the two sessions was greater than the difference perceived by the participants with more expertise ($F(1, 19) = 22.805, p<0.001$), with a large effect size – η^2_{partial} = 0.546. The effect of familiarity was approaching significance ($F(1, 19) = 3.281, p = 0.086$, η^2_{partial} = 0.147), and there was a significant interaction effect ($F(1, 19) = 11.864, p = 0.003$, η^2_{partial} = 0.384).

Playing the game twice could have potentially resulted in higher immersion scores in the second session, as people get more engaged in the process of playing, and possibly learn from the mistakes they make in the first session, although the tutorial session was aimed to try and account for the training

Component	Adaptive AI		Standard AI		F(1, 19)	p	η^2_{partial}
	Mean	Std.Dev.	Mean	Std.Dev.			
Total Immersion	119.86	9.55	111.90	14.29	7.88	0.011*	0.283
Cognitive Involvement	36.86	3.04	34.14	5.78	7.09	0.015*	0.262
Emotional Involvement	22.48	3.23	21.43	3.50	3.05	0.096	0.132
Real World Dissociation	26.05	3.71	23.95	3.64	5.84	0.025*	0.226
Challenge	13.81	2.32	12.76	2.17	2.41	0.136	0.108
Control	20.67	2.24	19.62	2.76	3.58	0.073	0.152

Table 1. Average levels of immersion and its components when playing the game with and without adaptive AI.

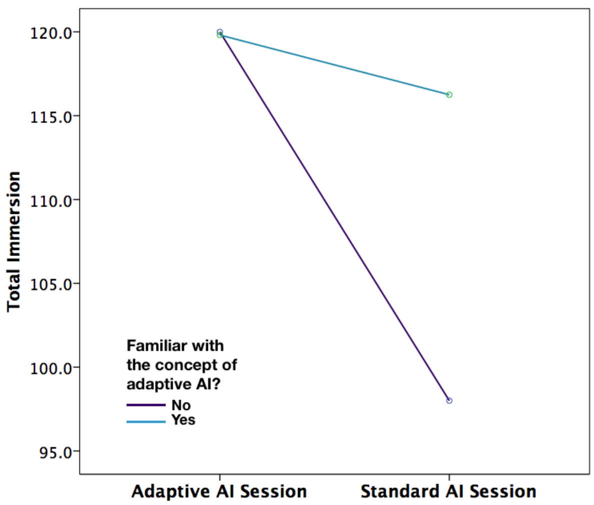

Figure 3. The effect of participants' familiarity with the concept of adaptive AI on total immersion in each section.

effect. However, the order of the sessions did not have any significant effect on the results ($F(1, 19) = 0.92, p = 0.350$, η^2_{partial} = 0.046) – 'adaptive' sessions were perceived as significantly more immersive regardless of the order of gaming rounds ($F(1, 19) = 7.58, p = 0.013$, η^2_{partial} = 0.285).

Quantitative Comparison of Two Sessions

The final questionnaire was used to collect data about participants perception of the adaptive AI in the game, and its effect on their enjoyment and performance. Ten participants answered that the game learned from and adapted to their behaviour, while five thought that the game did not adapt as such. When asked to rate the extent to which the game adapted to their behaviour as a player, participants gave it an average score of *3.2 (sd = 1.08)* , where 1 meant that the game did not adapt at all and 5 meant the AI was adjusting the game a lot according to the player's actions.

In terms of the overall enjoyment of the game, most participants stated that the adaptive AI made a whole lot of difference. They gave it an average rating of *3.7 (sd = 1.27)* out of 5, where 1 meant that the adaptive AI had no effect on their enjoyment, and major difference was rated as 5. When asked whether the gaming session with adaptive AI was more (a rating of 5) or less enjoyable (a rating of 1) than the randomly generated one, participants stated that the adaptive game made the playthrough more entertaining, giving it an average score of *3.9 (sd =1.14)*.

Similarly, participants believed that the 'presence' of adaptive AI was having an effect on their performance in the game. They gave the adaptive session an average rating of *3.5 (sd = 1.03)*, where 1 meant that this feature did not affect their performance and 5 meant that their performance was largely influenced by the presence of the AI. Participants were then asked to evaluate the impact of the adaptive AI on their performance – if a player thought that their performance was improved by the presence of the feature in the game, they gave it a rating of 1, but if they felt that they performed worse in this session, they gave it a rating of 5. In this question, the average rating was *2.6 (sd = 0.8)*, meaning that participants believed that the adaptive AI positively influenced their performance.

Eleven participants stated that the session with adaptive AI was more challenging than the standard one (a rating of 5), while the other 10 thought the opposite (a rating of 1), which resulted in an average rating of *3.2 (sd = 1.38)*. Nobody said that the sessions were equally challenging.

At the end of the questionnaire everybody provided a brief description of what adaptive AI did based on their experience and understanding. Following the end of the experiment, all participants were convinced that there was in fact an AI adapting to their play, although some stated that the quality of this adaptation could be improved. Every player was surprised to find out that both sessions were randomly generated, with no adaptive AI to affect their playthrough. However, two people said they were sceptical of the plausibility of having this feature in the game – both had more expertise due to their study in the area of AI and a more thorough understanding the mechanics behind the adaptation of game AI.

Qualitative Comparison of Two Sessions

It is evident that all participants experienced the adaptive AI in the particular session they were told to expect it in. People believed that the game was providing them with new objects in quantities they needed and in the locations they required them the most. Some participants found more useful objects: *"The adaptive AI put me into a safer environment and seemed to present me with resources as needed"*, *"I think the adaptive AI makes objects in the game appear more often when I need them. It reduces the time of exploring the map which makes the game more enjoyable"*. While some players assumed that their current inventory was affecting the objects they could locate: *"The adaptive AI seemed to be aware of the materials I needed to progress, and provided them with* *easy access. It also changed the number of monsters depending on how well-equipped I was"*, *"It seemed to move some of the things I collected a lot of further away, separating some of the elements of tools I built a lot."*

Players also believed that the monsters could adjust their behaviour depending on the things they have previously done in the game, e.g. *"Avoiding insect nests seemed to result in an abundance of them in newly explored areas. The first night a spider walked into my circle of light then ran away as I approached, as a result of me no following it out of my campfire area, the tactic seemed to change and on the second night a group of 3 spiders just charged up to my character"*. Many participants found the 'adaptive' session more challenging than the standard one, as they encountered more monsters: *"First thing I noticed is that there are far more dangers than the previous session. Second, they chased me for longer time. But in terms of their behaviours, I couldn't tell much difference"*, and discovered fewer objects they believed they needed in the game: *"[the adaptive AI was] trying to counteract my previous behaviour in game, i.e. prevent me from discovering too many things at once and more scattered around the map."*

However, due to the random factor that allowed for the generation of brand new worlds every time a player had a new session, it was possible to argue that one of the games was 'more adaptive' than the other. Although, technically, participants were rating the same game after each round, they experienced each round differently, and not only because of the randomly created objects and places. The players believed that one of the sessions was better than the other one in some way, sometimes more challenging than the other: *"To me it seemed like the difficulty level was increasing too rapidly. Even when I was not yet able to navigate the world, this AI produced monsters that then killed me before I could finish the task at hand (building fire?). For my own gaming style, the AI acted unfavourably. It should have supported my learning of the environment. Instead, it was there to make me adapt fast or die. That is more realistic, but also much more frustrating for a novice. AI therefore seems to me like a feature suitable for advanced players"*, sometimes making the playthrough easier: *"...Seemed to be more of what you needed nearby e.g. when low on health there were more flowers around."*

It is evident that regardless of the changes made by the random factor in the game, the players explained the differences in terms of the adaptive AI, i.e. they were actively seeking explanations during game play, and this as a result affected their experience. Therefore, independently of the immersion measures, providing players with clues of what to expect changed their interpretation of the experiences.

Discussion

The results support our hypothesis that players are able to influence their experience of a digital game based on a placebo effect. It is sufficient to say that there is a new technology in the game for the game to become more immersive even without the details of what the technology is or what it will do. Players are using their own beliefs and knowledge to generate the experience. If they choose to believe that a game has an adaptive AI, which supposedly can make their gaming session more enjoyable and enhance their performance, they themselves will subconsciously lead to these outcomes. Particularly in this situation, where an authority (an experiment facilitator) provided clear instructions, the players is particular susceptible to believing what is being said.

As well as the quantitative differences in immersion, the comments of participants were on par with the data collected in the IEQ. Some participants perceived the 'adaptive' session as more challenging, and some experienced the opposite according to their needs. This difference in the perception was due to randomness in generating objects and monsters, providing different levels of difficulty for each player, which resulted in no significant difference in terms of challenge. However, it is possible to draw a line between players' performances and their understanding of what the AI did. Players who felt challenged in one session, perceived the 'adaptive' session as trying to help them to perform better. Conversely, under-challenged players thought that the 'adaptive AI' was trying to increase the difficulty by adding new challenges, such as generating more enemies and less food.

From the components of immersion, it seems the increase is due to increased cognitive involvement and real world dissociation. This could be because players are seeing the game as providing a better experience or it may be that in trying to work out the adaptations, they are thinking harder about the game.

The level of expertise in the field also greatly affected the persuasiveness of each player. People tend to interpret events according to their own knowledge and heuristics. However, if they are given an explanation that is beyond their level of expertise, they would believe in what they hear or read, as long as the explanation is plausible. The events are then interpreted according to the new information received. It is therefore possible to experience something that is not there, like in the case of adaptive AI. As a result, participants who did not have any previous experience with adaptive AI were more likely to 'experience it', than those people who had extensive knowledge of the adaptive mechanisms.

This experiment is intended to reflect a common way in which players play, namely having played other versions or having other versions to compare to. However, there is an obvious risk of desirability bias in this experiment even though players were not talked what to expect from the 'adaptive AI.' This is equally present in some play-testing situations, but nonetheless is a threat to the validity of the findings of this experiment. We therefore aim to overcome this threat in the next study.

STUDY II: BETWEEN-SUBJECT DESIGN

The previous study was conducted in order to understand players' behaviour when comparing two gaming sessions, where one is supposedly able to adapt to their game play. However, because participants were exposed to two sessions and were asked to compare them, it is possible that players just looked for differences in the sessions and used them to describe their perception of adaptation in the game. We,

Component	Adaptive AI		Standard AI		F(1,38)	p	η^2_{partial}
	Mean	Std.Dev.	Mean	Std.Dev.			
Total Immersion	124.15	6.51	116.00	13.37	6.01	0.019*	0.137
Cognitive Involvement	38.25	2.65	35.75	5.02	3.87	0.056	0.092
Emotional Involvement	23.05	2.82	21.20	3.46	3.44	0.071	0.083
Real World Dissociation	27.55	3.52	24.85	4.86	4.05	0.051	0.096
Challenge	14.20	1.74	14.10	2.10	0.03	0.870	0.001
Control	21.10	1.74	20.10	2.90	1.75	0.194	0.044

Table 2. Immersion and its components in control and experimental groups.

therefore, conducted another study in order to test whether this was indeed the case, and to explore players' perception and experiences of a digital game with suggested feature based on a single impression.

This experiment is therefore *ceteris paribus* a between-subject version of the previous one. There is one dependent variable, immersion, measured using the IEQ. The experimental manipulation is the same as before, whether participants are told that the game has adaptive AI or not.

Experimental Method

Overall, 40 participants (9 women and 31 men) took part in the study. The recruited people were students at the University of York from diverse backgrounds, and with varied levels of gaming experience. The average age of the participants was 23.5 ($sd = 6.32$), with the age range between 18 and 43 years.

Most participants regularly played digital games, often spending over an hour in a single session, several times a week. Those players listed strategies, adventure games, RPG and FPS games as the genres they prefer; while puzzle, action and sports games were also amongst some of the less frequently mentioned ones. There were four people who stated that they do not often play games, but when they do – they normally spend an hour or more in a single gaming session playing puzzle games, sports, action and adventure games.

Out of 20 people in the experimental group, 12 stated that they were familiar with the concept of adaptive AI, but only three of those participants had knowingly played games with adaptable behaviour prior to the experiment.

Materials

The game used for this experiment was also 'Don't Starve'. None of the participants have previously played the game, even though some people had previously heard about it.

The PX data was collected using the IEQ. An additional questionnaire was used to measure the players' perceived extent to which the game was able to adapt to their behaviour – the questionnaire was deliberately phrased in a way that players in both groups could evaluate the responsiveness and the adaptiveness of the game with or without prior knowledge about the game having or not having adaptive AI, depending on the group. Additionally for the participants in the experimental group, the questionnaire was completed with an open-ended question, asking them to describe what they thought the adaptive AI did in this game based on their evaluation.

Procedure

The experiment started with a demographics questionnaire to gather the data about participants' gaming background. This was followed by a short tutorial, during which participants were provided with a brief explanation of the aim, the controls and the storyline of the chosen game. This was then followed by their first game trial, when they played for about 5 minutes to make themselves familiar with the virtual environment. For the main part of the study, participants played the game for 20 more minutes, starting from the beginning, without the experiment facilitator present in the room.

All participants in the experimental group received an explanation of what adaptive AI is and does before they played the game on their own – the same explanation as used before for the within-subject study. Each participant played the game from the very beginning, and each time the world was randomly generated with settings being at default values for both groups of participants. At the end of the gaming session they filled in the IEQ, and the questionnaire with eight Likert scale questions designed to collect data about players' perception of the 'adaptiveness' of the game.

Results

The findings were on par with the results obtained in the previous study. The hypothesis that the participants playing the game 'with adaptive AI' would feel immersed in the game more than the control group was confirmed. On average, immersion scores collected in the control group were lower than the scores obtained in the experimental group, with a significant difference between these two groups as determined by one-way ANOVA ($F(1, 38) = 6.01, p = 0.019$), and a small effect size ($\eta^2_{\text{partial}} = 0.137$) (Figure 4). Participants who were informed about the game 'having adaptive AI' had an average score of 124.15 ($sd = 6.51$), whereas people who played the game without being told anything about the feature had a mean score of 116.00, with a larger variation in the results ($sd = 13.37$). Standard deviations are within acceptable range for assuming homogeneous variances [3].

When comparing the immersion components separately, the difference in scores of three components was approaching significance, while two did not have significant difference in the results (table 2). The difference in the scores for cognitive involvement and real-world dissociation was approaching significance. Moreover, participants in the control group felt less emotionally attached to the game than the experimental group, with the difference approaching significance. There was also no significant difference between the two groups in terms of the perceived challenge and the control mastery.

Questions	Standard AI		Adaptive AI		F(1, 38)	p	η^2_{partial}
	Mean	Std.Dev.	Mean	Std.Dev.			
The game was generating content according to my behaviour in the game.	2.55	1.32	3.40	0.88	5.75	0.022*	0.131
New content in the game appeared based on my decisions as a player.	2.55	1.23	3.45	0.76	7.71	0.008**	0.169
The game matched the challenge to my skills and abilities as a player.	2.50	1.05	3.50	0.95	10.00	0.003**	0.208
The behaviour of the game changed when I was doing too well or too badly.	1.90	1.02	3.20	1.06	15.66	0.000***	0.208
The game was generating contents based on the needs of my character at that point in the game.	3.05	0.95	3.90	0.85	8.93	0.005**	0.190
The game was not responding sensibly to my actions as a player.	2.15	0.88	1.90	0.97	0.73	0.397	0.019

Table 3. Average agreeableness with the statements for participants in different groups based on their gaming experience.

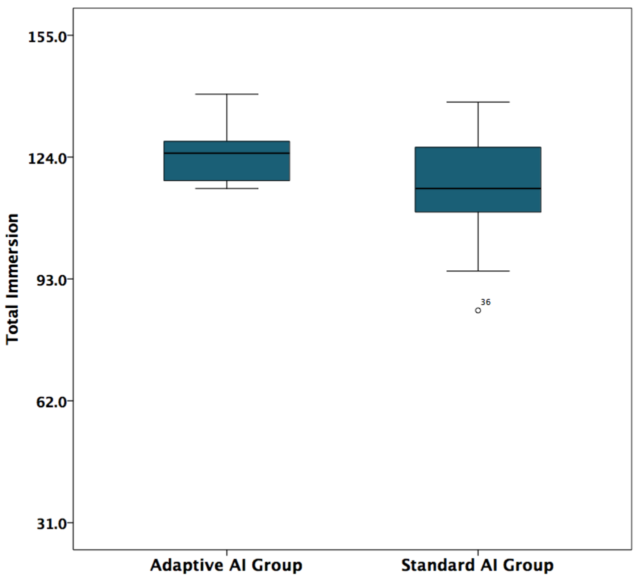

Figure 4. Total immersion with and without adaptive AI.

Familiarity with the concept of adaptive AI had a marginally significant effect on immersion scores in the experimental group: $F(1, 19) = 2.93, p = 0.066$ ($\eta^2_{\text{partial}} = 0.137$), where 12 out of 20 participants had had some knowledge about adaptive AI before the experiment.

Quantitative Evaluation of Adaptiveness

A Likert-scale questionnaire at the end of the session was aimed at collecting data about players' perception of the game's adaptiveness. When answering questions, each participant chose a number from 1 to 5 based on how much they agreed with the statement. Table 3 shows 6 out of 8 questions, which each group of participants were asked after filling in the IEQ, as well as the average scores, provided by each group of participants.

Overall, the results demonstrate that participants in the experimental group felt that the game was altering its behaviour based on their decisions more than those players, who were not aware of this adaptive feature.

Additionally, players in the control group on average evaluated their performance in the game lower: $\overline{X}_{\text{C}} = 3.15$ ($sd = 0.93$) than those participants, who were told about presence of adaptive AI: $\overline{X}_{\text{E}} = 3.35$ ($sd = 0.86$), however this difference was not significant: $F(1, 38) = 0.49, p = 0.489$ ($\eta^2_{\text{partial}} = 0.013$).

Similarly, players in the experimental group also believed that they enjoyed the game more: $\overline{X}_{\text{E}} = 4.25$ ($sd = 0.64$) than the control group: $\overline{X}_{\text{C}} = 4.60$ ($sd = 0.60$), with the difference between the two approaching significance: $F(1, 38) = 3.20, p = 0.082$ ($\eta^2_{\text{partial}} = 0.078$).

Qualitative Evaluation of Adaptiveness

Additional qualitative data was collected in the experimental group in order to understand whether the players were able to perceive the adaptive AI, and if they did, what exactly players though it was doing.

Generally, participants believed that the game was adapting to their behaviour in the game, however as they did not have anything to compare their experience to, the answers provided by the players were more vague than the answers collected from participants in within-subject design study.

The majority of participants thought that the game was adapting to them in a positive manner. The main adaptations that they spotted were normally related to food and monster locations and quantities, based on their character's needs at a specific point in the game: *"I think it was disturbing material I needed, but not to a extend that they were easily gained. The difficulty of the environment and the animals I encountered changed the longer I played."* and *"[the game was] generating objects needed for my progress in the game based on the previous selection of tools and movements.", "... It also chose appropriate enemy strength for me."*

Participants also found that the game was adapting in order to keep them in suspense, and provide new challenges and entertainments: *"Introducing the players to the environment at a high difficulty level then altering the environment depending on how well the player responded. I also think it provided materials suited to the players needs quite well."* and *"Trying to make me use the parts of the game I wasn't using, for example I was not attacking things so after a while I was attacked.*

It was trying to make the game more challenging by introducing more obstacles., or more enemies: *"More than anything, it felt like the game was adding more enemies. In some areas with lots of them there seemed to be more/better rewards to get me to fight/interact with the enemies to get them."*. Some players even thought that the game was being a little mean to them: *"I think it was avoiding me with the things I needed most - i.e. food, wood (basics) so that I could progress at a steady level whilst withholding other materials (stone) so that I did not progress too quickly."*

However, not all participants were entirely convinced that there was any adaptation happening, or if it was, it was not being particularly responsive to their game play: *"I felt the adaptive AI was very subtle..."*. Some players pointed out that playing the game only once meant that they did not have any other experiences to compare to: *"The AI could have been giving me appropriate objects and enemies, but it felt like objects were random and its possible that I just didn't meet the enemies. I have no expectations - it's a random environment so difficult to tell if it's changed."*

Discussion

The results of this study confirmed the hypothesis and were very similar to the ones obtained in the first study. The findings support the idea that PX, specifically immersion, is formed not only based on the features in a game but also on players' personal understanding of how the game should work regardless of whether that understanding is correct. The results demonstrate that players' expectations of a digital game can be adjusted based on their knowledge of the game before playing it, which then subsequently leads to an improved experience if the player chooses to believe this information.

Interestingly, the differences in the components of immersion were on par with the finding in the previous study. The experimental group were more cognitively involved, which led to a more focused state, dissociating the player from the real world. The effect size between the two groups was not as strongly significant as it was seen in the previous study, but this is to be expected because players in this study are not acting as their own controls, this being a between-subjects design. This does suggest that players were becoming genuinely more involved in the game as a result of the manipulation even though the perceived challenge was the same.

The additional questionnaire helped to understand whether participants were able to perceive the game's 'adaptiveness'. The answers given by players in the experimental group were significantly different from the control group's responses. Knowing about adaptive AI made players believe that the game was changing its behaviour based on their performance and their decisions, and generating appropriate content for the needs of their character. It is worth noting that without the other version of the game to compare to, the comments seemed to lack the same level of confidence of the previous study. This contrasts with the control group, who perceived the game as it was, namely, placing objects and non-player characters randomly around the map.

Previous knowledge of adaptive AI also had an effect on the perception of this feature. Players with more expert knowledge in the field were more sceptical about the adaptation used in the game than those participants, who had not heard the term before taking part in the study.

CONCLUSIONS

Both studies confirmed that it is possible to implant an idea into players' minds that a digital game is capable of performing something that it is not able to do. In this particular situation players are then able to experience this feature that is supposedly beneficial for their experience depending on the plausibility of explanation. The players adjust their expectations of the game based on these suggestions, which then affects their experience of the game.

The intention of these experiments was not to prove that due to players' extensive imagination digital games do not need improvement, but instead the idea was to demonstrate that the experience that comes from playing games can be a result of gamers' personal attitude and expectations.

This work therefore has two important implications. First, any experimental investigation into the influence of new features in a game, such as adaptive AI, on PX must be made carefully, without any opportunity for players to second guess what the investigation is about. For example, a study in an AI lab that uses an existing game may trigger the expectation of 'something good.' The mere expectation of a difference can be sufficient to change the experience. Ideally, further experiments in this area will need to be conducted like a double-blind randomised control trial where neither experimenter nor participant know the experimental manipulation involved. This becomes even more challenging in the context of play-testing where surely players called in for play-testing must be expecting something new even if it is not explicitly communicated what.

Secondly, given the prevalence of sequels in the game industry, we must take any claims for advances in the underlying technology with a pinch of salt. Players may have an improved experience over earlier versions of the game simply because they expect it. Evaluating the real effects of new versions of a game will need to be done very cautiously if companies are not to end believing the claims of their own hype (mediated by their willing play-testers).

It may be of course that such effects seen here wear off over time so that over a longer timescale, the 'true' experiential outcomes are achieved but this is currently unknown. It may be that players work out quite quickly the lack of advanced technology or it may be that they persist in their misconceptions thanks to the careful explanations that they provide themselves. It is also unclear just how much information a player needs to be susceptible to the effects seen here. Different technologies also need to be tested in order to be able to explore whether these findings hold true for other kinds of digital games. These are all important avenues for further exploration.

REFERENCES

1. Bartle, R. Hearts, clubs, diamonds, spades: Players who suit muds. *Journal of MUD research 1*, 1 (1996), 19.
2. Blackwell, B., Bloomfield, S. S., and Buncher, C. R. Demonstration to medical students of placebo responses and non-drug factors. *The Lancet 299*, 7763 (1972), 1279–1282.
3. Box, G. E., et al. Some theorems on quadratic forms applied in the study of analysis of variance problems, i. effect of inequality of variance in the one-way classification. *The annals of mathematical statistics 25*, 2 (1954), 290–302.
4. Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., and Pidruzny, J. N. The development of the game engagement questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology 45*, 4 (2009), 624–634.
5. Brown, E., and Cairns, P. A grounded investigation of game immersion. In *CHI'04 extended abstracts on Human factors in computing systems*, ACM (2004), 1297–1300.
6. Cairns, P., Cox, A., and Nordin, A. I. *Immersion in Digital Games: Review of Gaming Experience Research*. John Wiley & Sons, Inc., 2014, 337–361.
7. Charles, D., Kerr, A., McNeill, M., McAlister, M., Black, M., Kücklich, J., Moore, A., and Stringer, K. Player-centred game design: Player modelling and adaptive digital games. In *Proceedings of the Digital Games Research Conference*, vol. 285 (2005).
8. Cox, A., Cairns, P., Shah, P., and Carroll, M. Not doing but thinking: the role of challenge in the gaming experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2012), 79–88.
9. Csikszentmihalyi, M. *Flow: The psychology of optimal experience*. New York: Harper Perennial, 1991.
10. de Craen, A. J., Tijssen, J., de Gans, J., and Kleijnen, J. Placebo effect in the acute treatment of migraine: subcutaneous placebos are better than oral placebos. *Journal of neurology 247*, 3 (2000), 183–188.
11. Geers, A. L., Weiland, P. E., Kosbab, K., Landry, S. J., and Helfer, S. G. Goal activation, expectations, and the placebo effect. *Journal of personality and social psychology 89*, 2 (2005), 143.
12. Goldacre, B. *Bad science: quacks, hacks, and big pharma flacks*. McClelland & Stewart, 2010.
13. Huhtala, J., Isokoski, P., and Ovaska, S. The usefulness of an immersion questionnaire in game development. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, ACM (2012), 1859–1864.
14. Jenkins, R., Lee, M., Archambault, R., Shane, H., and Greg, S. The influence of professional critic reviews. *EEDAR/SMU behavioral study. Tech. rep., Southern Methodist University, Guildhall. Electronic Entertainment Design and Research* (2010).
15. Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., and Walton, A. Measuring and defining the experience of immersion in games. *International journal of human-computer studies 66*, 9 (2008), 641–661.
16. Johnson, S. Adaptive AI: A practical example. *AI Game Programming Wisdom 2* (2004), 639–647.
17. Kerr, A. Girls/women just want to have fun-a study of adult female players of digital games. *Paper presented at the Level up. Digital Games Research Conference Utrecht, The Netherlands* (2003).
18. Kikas, E. University students' conceptions of different physical phenomena. *Journal of Adult Development 10*, 3 (2003), 139–150.
19. Klei Entertainment. Don't Starve. `http://www.dontstarvegame.com/`, 2013.
20. Livingston, I. J., Nacke, L. E., and Mandryk, R. L. Influencing experience: the effects of reading game reviews on player experience. In *Entertainment Computing–ICEC 2011*. Springer, 2011, 89–100.
21. McMahan, A. Immersion, engagement and presence: A method for analyzing 3-d video games. In *The video game theory reader*, M. Wolf and B. Perron, Eds. Routledge, London & New York, 2003, 67–86.
22. Nordin, A. I., Cairns, P., Hudson, M., Alonso, A., and Calvillo, E. H. The effect of surroundings on gaming experience. *Foundations of Digital Games* (2014).
23. Rigby, S., and Ryan, R. M. *Glued to Games: How Video Games Draw Us In and Hold Us Spellbound: How Video Games Draw Us In and Hold Us Spellbound*. ABC-CLIO, 2011.
24. Sanders, T., and Cairns, P. Time perception, immersion and music in videogames. In *Proceedings of the 24th BCS Interaction Specialist Group Conference*, British Computer Society (2010), 160–167.
25. Schell, J. *The Art of Game Design: A book of lenses*. CRC Press, 2014.
26. Sundar, S. S., Knobloch-Westerwick, S., and Hastall, M. R. News cues: Information scent and cognitive heuristics. *Journal of the American Society for Information Science and Technology 58*, 3 (2007), 366–378.
27. Sundar, S. S., Xu, Q., and Oeldorf-Hirsch, A. Authority vs. peer: how interface cues influence users. In *CHI'09 Extended Abstracts on Human Factors in Computing Systems*, ACM (2009), 4231–4236.
28. Thompson, M., Nordin, A. I., and Cairns, P. Effect of touch-screen size on game immersion. In *Proceedings of the 26th Annual BCS Interaction Specialist Group Conference on People and Computers*, British Computer Society (2012), 280–285.

29. Weisberg, D. S., Keil, F. C., Goodstein, J., Rawson, E., and Gray, J. R. The seductive allure of neuroscience explanations. *Journal of cognitive neuroscience 20*, 3 (2008), 470–477.

30. Wirth, W., Hartmann, T., Böcking, S., Vorderer, P., Klimmt, C., Schramm, H., Saari, T., Laarni, J., Ravaja, N., Gouveia, F. R., et al. A process model of the formation of spatial presence experiences. *Media psychology 9*, 3 (2007), 493–525.