The Convergence of Player Experience Questionnaires

Alena Denisova
Department of Computer Science, University of York
York, United Kingdom
ad595@york.ac.uk

A. Imran Nordin
Institute of Visual Informatics, The National University of Malaysia, UKM Bangi
Selangor, Malaysia
imran.nordin@gmail.com

Paul Cairns
Department of Computer Science, University of York
York, United Kingdom
paul.cairns@york.ac.uk

ABSTRACT
Player experience is an important field of digital games research to understand how games influence players. A common way to directly measure players' reported experiences is through questionnaires. However, the large number of questionnaires currently in use introduces several challenges both in terms of selecting suitable measures and comparing results across studies. In this paper, we review some of the most widely known and used questionnaires and focus on the immersive experience questionnaire (IEQ), the game engagement questionnaire (GEQ), and the player experience of need satisfaction (PENS), with the aim to position each of them in relation to each other. This was done through an online survey, in which we gathered 270 responses from players about their most recent experience of a digital game. Our findings show considerable convergence between these three questionnaires and that there is room to refine them into a more widely applicable measure of general game engagement.

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

Author Keywords
Player experience; immersion; engagement; player experience of need satisfaction; questionnaires; PENS; IEQ; GEQ.

INTRODUCTION
Flow, presence, engagement, immersion, and fun are amongst most commonly used terms to describe the experience people have when playing digital games [2]. Many methods to evaluate these experiences exist. These include objective measures, such as heart rate measurements, electromyography (EMG) and electrodermal activity (EDA), but these are hard to map to the subjective experience of players. More appropriate subjective measures are also used, such as interviews, focus groups, and surveys, but these can lack standardisation and comparability. Questionnaires are useful standardised research instruments that allow quantification of the subjective experience under consideration, while being relatively easy to deploy [1]. Like the more objective measures, the use of questionnaires ensures consistency and uniformity of collected data, because the same specific aspects are considered by all participants in all studies.

There are, however, a few drawbacks of using questionnaires to measure player experience. Nordin et al. [12] named the challenges researchers face when looking for the most appropriate questionnaire. Amongst these, they note, is the ability to persuade participants to treat the questionnaires seriously, and the scale upon which participants answer them. Moreover, it is important to consider the wording of questions, so it does not reduce the face validity of the questionnaires [1].

Player experience is a multi-faceted experience. Theories, and their corresponding questionnaires, aim to address each unique concept in great detail. While some questionnaires measure broader experiences, such as engagement and immersion in games [3, 11], which take into account most aspects of gaming, others focus more on a specific facet of experience, e.g. narrative immersion or social presence [7, 13].

On the one hand, the variety of questionnaires allows researchers to focus on a specific aspect of games. At the same time, the various questionnaires show considerable conceptual, and in some cases actual, overlap, while supposedly measuring apparently different experiences. This leads to a confusion as to whether they in fact do the same job. The plurality of questionnaires also reduces the ability to compare the outcomes of player experience studies. The aim of this work is to see empirically how three of the most widely used questionnaires conceptually converge or diverge. The goal is not to say that there should be only one questionnaire, but instead to evaluate whether questionnaires with similar aims to measure engagement in digital games produce consistent and correlated results. This empirical work also helps to determine which aspects of these questionnaires work well and which do not. As a result, we also suggest improvements to the extant tools for quantifying a general game engagement.

MEASURING PLAYER EXPERIENCE
Many existing player experience theories use their own questionnaires to quantify the experience one is having when playing digital games. While the theories aim to focus on a specific aspect of player experience unique to each concept, the overlap between the theories is evident. Similarly, the measuring tools each theory uses have much in common in their questions and

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
CHI PLAY '16, October 16 - 19, 2016, Austin, TX, USA
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-4456-2/16/10 $15.00
DOI: http://dx.doi.org/10.1145/2967934.2968095

Questionnaire	*Components*
Immersive Experience Questionnaire (IEQ)[11]	Cognitive Involvement, Emotional Involvement, Real World Dissociation, Challenge, Control
Game Engagement Questionnaire (GEQ/GEngQ)[3]	Absorption, Flow, Presence, Immersion
Player Experience of Need Satisfaction (PENS)[14]	Competence, Autonomy, Relatedness, Controls, Presence/Immersion

Table 1. Questionnaires measuring player experience after playing a digital game.

components, as seen in Table 1, which contains a summary of the three widely known and used questionnaires and their components. These three questionnaires aim to measure immersion either as a component of player experience, like in the case of the GEQ and the PENS, or as a whole experience, as the IEQ does.

In addition to these three questionnaires, some examples of other important questionnaires used to measure player experience include the Game Experience Questionnaire (GExpQ) [10], GameFlow [16], the questionnaire used to measure immersion by Ermi and Mäyrä [9], as well as more broad concepts like Flow [6], and Presence [17] questionnaires. Such a large number of existing questionnaires poses a challenge for new researchers, who may not necessarily be familiar with every specific detail of each theory. Choosing one is therefore often based on their availability: questionnaires like the GExpQ or the immersion questionnaire by Ermi and Mäyrä are not readily available to the researchers. So eventually only the easily accessible questionnaires tend to be used for measuring player experience. There is also a question of reliability. To obtain reliable results it is imperative that the data is gathered using a reliable questionnaire. However, some of the available questionnaires are not statistically validated, and as a result cannot be presumed trustworthy.

We, therefore, decided to focus on the three tools outlined above based on their dominant use in gaming research, their availability, and the conceptual overlap. The GEQ (GEngQ) [3] and the IEQ [11] are both available publicly and are set up in a similar fashion to evaluate player experience. The GEQ was developed to assess the deep engagement of violent video game players, and it consists of 19 positively worded questions answered on a 7-point Likert scale. The questionnaire is formulated in such a way that the engagement is a unidimensional experience, which ranges up from immersion to flow.

The IEQ is used to measure the levels of immersion experienced by players. It has been used extensively across a diverse array of different use cases and game genres, for example [5, 8]. The IEQ uses 5-point Likert scale questions to measure player experience, but is specifically focused on the notion of immersion when playing games. It uses a combination of positively and negatively worded statements, adding an additional layer of accuracy. The overall score is composed of a summary of the results from the positive questions, and the inverted results of the negative. The development of the IEQ also suggested that there are five factors underlying immersion, but in practice, immersion is treated as a unidimensional concept with the factors framing the interpretation of the results.

Another questionnaire frequently used to quantify the experience of playing digital games is the player experience of need satisfaction (PENS). The questionnaire contains 19 items, where it reviews the experience in terms of 5 components, such as competence, autonomy, relatedness, immersion/presence, and intuitive controls. All but one are measured using 3-item scales (apart from immersion, which is a 9-item scale), ranked on a 7-point Likert scale. It has been statistically validated [15], however the questionnaire is copyrighted and therefore is not readily available to researchers.

An item-by-item analysis shows some similarities between all three of these questionnaires. It is reasonable to expect some correlation between the results obtained using them. However, all three are also described as measuring differing concepts, with the PENS in particular addressing five ostensibly unrelated aspects of player experience. The question is to what extent these questionnaires do in fact measure different concepts.

EXPERIMENTAL METHOD

The aim of the present study was to compare three of the most widely used questionnaires measuring player experience: the IEQ, the GEQ, and the PENS. For this, their questions were combined into an online survey, which was distributed in a number of online gaming forums in order to gather responses from a variety of digital game players.

Participants

Overall, the study gained 287 respondents, where 17 entities had to be omitted due to the age restrictions imposed by the ethical clearance of the study, leaving 270 valid responses from 30 women, 232 men, 1 person who identified themselves as other gender, and 3 people who did not report their gender. Their average age was 26.42 years ($SD = 6.66$, min/max : $18/63$). Participants were mostly native English speakers, and were from a total of 32 countries. They had a varied level of previous experience of playing digital games, averaging 17.5 years of gaming ($SD = 6.63$).

Participants were invited to complete the survey, in which they had to reflect on their most recent experience of playing a digital game, which they entered before taking the survey. Overall, over 100 titles were entered, with some of the most popular games being “The Witcher 3” (13), “Dark Souls” series (8), “League of Legends” (4), “FIFA” series (8), “DOTA 2” (19), “Fallout 4” (9), “Counter-Strike” series (7) and “Enter the Gungeon” (9). Other titles listed were from a variety of genres, including role-playing games (RPGs), action games, and action-adventure games of various kinds, simulations, strategy,

		IEQ	GEQ	PENS (Total)	PENS				
					Competence	Autonomy	Relatedness	Immersion	Controls
		$M = 141.56$ $SD = 22.71$	$M = 67.90$ $SD = 16.61$	$M = 94.84$ $SD = 20.46$	$M = 15.40$ $SD = 3.46$	$M = 15.60$ $SD = 3.83$	$M = 11.61$ $SD = 4.30$	$M = 36.61$ $SD = 11.86$	$M = 15.62$ $SD = 3.85$
	IEQ	–							
	GEQ	0.804**	–						
	PENS (Total)	0.813**	0.692**	–					
PENS	**Competence**	0.573**	0.405**	0.592**	–				
	Autonomy	0.595**	0.428**	0.697**	0.443**	–			
	Relatedness	0.461**	0.421**	0.683**	0.237**	0.333**	–		
	Immersion	0.705**	0.666**	0.902**	0.323**	0.500**	0.586**	–	
	Controls	0.524**	0.369**	0.546**	0.547**	0.399**	0.163**	0.270**	–

Table 2. Pearson r correlations of questionnaire scores ($N = 270$, $^{**}p < 0.01$).

and racing games. To incentivise the participants we offered them to be entered into a prize draw raffle to win Steam or Amazon vouchers worth £20, depending on their preference.

Materials

The questions from the IEQ, GEQ, and PENS questionnaires were merged to produce a single unified questionnaire that was delivered through Google Forms. Because each questionnaire had different question formats that might confuse participants, the items from all three were presented as standard Likert-type statements in the present tense (as in the GEQ). For example, a question in the original version of the IEQ reading as: "To what extent did you find the game easy?" was rephrased to: "I find the game easy" to match the conventions of the other two questionnaires. All items had a 7-point Likert scale anchored at the ends with Strongly Disagree and Strongly Agree. The order of the questions was randomised in Google forms for each participant in order to avoid order-effects.

At the end of the questionnaire, there was an open-ended field for comments. This was not extensively used but, where appropriate, these responses are reported on.

Design and Procedure

The link to the survey was distributed on various online forums, such as the Steam Users' Forum, Twitter, and relevant Facebook groups, with the aim to gather responses from a diverse audience of digital game players. Each participant was briefed on the usage of the data in accordance with the ethical clearance provided on the study. After this they were asked to reflect back on their most recent experience of playing a digital game and to choose answers that best reflected their experience.

RESULTS AND DISCUSSION

Scale reliability was performed to ensure internal consistency for each questionnaire using Cronbach's α. Additionally, item-total correlations were considered to identify items with weaker coherence to the overall scales they belonged to. These are not reported here for brevity. Correlations between scales and their components were all calculated using Pearson's product correlations.

Scale Reliability and Principle Component Analysis

The collected data was used to perform reliability analyses on the questionnaires: the IEQ, the GEQ, PENS Competence, Autonomy, Relatedness, Immersion, and Controls scales. Internal consistency measures of reliability (Cronbach's α) for the IEQ and the GEQ yielded high levels of internal consistencies of 0.91 and 0.85, respectively.

Each of the PENS factors also had high levels of reliability: 0.74 for competence, 0.78 for autonomy, 0.88 for immersion, 0.80 for intuitive controls. However, relatedness had a lower internal consistency of 0.62, which can be considerably improved to 0.81 if one of the items is removed. Additionally, internal consistency of the PENS as a single scale was evaluated, yielding alpha of 0.90.

The PENS was not designed to be a uni-dimensional scale, and we recognise that high alpha is not a valid indicator of unidimensionality. We, therefore, conducted the Principal Component Analysis (PCA) on the 21 items with oblique rotation (direct oblimin). Analysis of the Measure of Sampling Adequacy suggested that the weak Relatedness item seen previously was not suitable for PCA and it was removed from further analysis. The scree plot strongly suggested two factors accounting for more than 50% variance and the structure matrix also suggested two clear factors, the first composed of Immersion and Relatedness ($\alpha = 0.90$) and the second of Autonomy, Competence, and Control ($\alpha = 0.84$). Details of the analysis are on our website[1].

This suggests that across the wide range of games considered by our participants, PENS does not automatically divide into five clear factors, but in this context has only two factors. It would be worth more substantially exploring the PENS to better understand why the conceptual differences underlying the scales are not seen in the PENS scores here.

Scale Correlations

Overall, there were high positive correlations between the pairs of the IEQ, the GEQ, and the PENS Immersion scales, as shown in the Table 2. The results obtained using the IEQ and the GEQ scales were highly correlated: 0.804. Similarly, the IEQ and the GEQ were also positively significantly correlated with the results from the PENS Immersion/Presence: 0.705

[1] sites.google.com/a/york.ac.uk/questionnaires/

and 0.666, respectively. The high correlation between the IEQ and the GEQ suggests that engagement and immersion are in fact addressing the same underlying aspect of player experience. Similarly, data gathered using the immersion scale of the PENS questionnaire also greatly correlated with results obtained using the other two. These findings are not surprising considering that engagement is often perceived as a part of immersive experience [4], and all three scales had questions of a similar nature.

Similarly, there was a positive significant correlation between players' perception of competence and to what extent they found the controls intuitive, according to the results collected using the PENS. As competence questions concerned players' perceived level of skill and challenge in the game, and controls questions were more relevant to the challenge players face when using the controls, it is fair to assume that there is a correlation between the two factors as they broadly address challenge, regardless of its nature. Having appropriate levels of challenge is important for the players to have a positive gaming experience, as reflected in the correlation between the competence and controls data and the IEQ results.

Autonomy, as it is measured in the PENS questionnaire, is described in terms of the amount of freedom and the interesting options the game offers their players. These questions were similar to the emotional involvement, as it is measured in the IEQ. Having interesting choices in the game also contributes to the overall experience, as it is also seen in the high correlation between autonomy and the IEQ and GEQ results. Similarly, there was a positive correlation between autonomy and relatedness. The two come hand in hand in games that offer opportunities for emotional involvement, and a storyline, in which the player can develop relationships with other players.

Given the high statistical reliability of the overall PENS scale, this was also treated as a single scale and compared to the other questionnaires and showed correlations of 0.813 and 0.692 with the IEQ and the GEQ, respectively. Interestingly, even the total scores of the items of the PENS scale that are not part of the Immersion component also correlated with the IEQ and GEQ, $r = 0.750$ and $r = 0.569$, respectively. This suggests it too is measuring engagement. From consideration of the questionnaire items, this is not so surprising. There is a large overlap between the themes of questions used in all three questionnaires, which address such aspects as physical and mental challenge, intuitive controls, emotional involvement (including relationships with other players, the storyline and aesthetics), sense of time, and a sense of being in the game world.

Problematic Items

Although the questionnaires produced coherent results, there were a few drawbacks. Some unreliable items became evident during the analysis of the collected data, such as questions asking players about their relationships with other players (PENS). As many single-player games do not provide opportunities for this experience, this question was viewed as confusing, and players left comments such as *(P13)*: "Some of the questions, for example the ones asking about my relationship to other player, didn't apply to a lot of the game (single player) games I prefer playing" and *(P99)*: "The questions that you were asking seemed to target more of a triple A game audience with world building or even more aptly a MMO or MMORPG I find those games to focus far more on relationship building, immersion and blurring the lines between reality and fantasy." As a whole however, relatedness to the "others was not always inapplicable, as many RPG games offer players opportunities to build relationships with other characters that can be valuable to the player. This perhaps suggests why one of the PENS relatedness questions did not function as well in the Relatedness scale where the other two did.

Another issue mentioned in the comments was about the fact that not all games have a clear ending, as one of the IEQ items concerns players' desire to "win" the game. Similarly, not all digital games are aimed at eliciting emotional responses, and therefore some items in the IEQ and the PENS were deemed inappropriate. A League of Legends player described his experience as something more akin to a sports player during a football match (*P241*): "The appeal of it isn't like Skyrim or The Witcher in the sense I want to be immersed in another world but more of the sense you get when you play a sport. I will be with a group of friends and out go "out" to play to forget about our worries and responsibilities bonding at the same time."

Moreover, some of the questions seemed out of place to some respondents, such as "I feel scared" item in the GEQ, which does not necessarily apply to many games. Vaguely phrased questions, such as "I feel different" (GEQ) also provided too many opportunities for interpretation, as well as the following item: "Things seem to happen automatically" (GEQ). These items also did not have strong correlations with the others.

CONCLUSION

Overall, the analysis of the collected data suggests that although there is much correlation between the three widely used questionnaires, there is potential for improvement. As different game genres elicit different aspects of gaming experience, the questionnaires in their present form are not fully applicable to all kinds of digital games. As things currently stand, all three seem to function as reasonable measures of player engagement in a game. However, we suggest that there is the opportunity to develop a more refined questionnaire based on these three, which is both a good measure of engagement and not dependent on the game or game type being played. This would not only allow more robust findings, but increase the comparability of studies in different contexts. There is some room to consider nuanced differences between aspects of engagement, for example, through relatedness or challenge but the IEQ and GEQ are not currently construed as addressing those nuances. In conclusion, in their present form, the questionnaires can be used equally reliably to measure player engagement generally. However, we argue that there should be a unified method, which allows us to evaluate players' experience in a variety of digital games without discriminating against games that do not have all such aspects. To do so, more research is needed to unveil the individual differences in games based on the theme, content, and styles of play.

REFERENCES

1. A. Adams and A. Cox. 2008. Questionnaires, in-depth interviews and focus groups. In *Research methods for human-computer interaction*, P. Cairns and A. Cox (Eds.). Cambridge University Press, Cambridge, UK, 17–34.
2. R. Bernhaupt, M. Eckschlager, and M. Tscheligi. 2007. Methods for evaluating games: how to measure usability and user experience in games?. In *Proceedings of the international conference on Advances in computer entertainment technology*. ACM, 309–310.
3. J. Brockmyer, C. Fox, K. Curtiss, E. McBroom, K. Burkhart, and J. Pidruzny. 2009. 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.
4. E. Brown and P. Cairns. 2004. A grounded investigation of game immersion. In *CHI '04 extended abstracts on Human factors in computing systems*. 1297–1300.
5. A. Cox, P. Cairns, P. Shah, and M. Carroll. 2012. Not doing but thinking: the role of challenge in the gaming experience. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*. ACM, 79–88.
6. M. Csikszentmihalyi. 1998. *The flow experience and its significance for human psychology*. Cambridge University Press, Cambridge, UK, Chapter In: Optimal experience: psychological studies of flow in consciousness, 15–35.
7. Y. De Kort, W. A IJsselsteijn, and K. Poels. 2007. Digital games as social presence technology: Development of the Social Presence in Gaming Questionnaire (SPGQ). *Proceedings of PRESENCE* 195203 (2007).
8. A. Denisova and P. Cairns. 2015. First Person vs. Third Person Perspective in Digital Games: Do Player Preferences Affect Immersion?. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 145–148.
9. L. Ermi and F. Mäyrä. 2005. Fundamental components of the gameplay experience: Analysing immersion. *Worlds in play: International perspectives on digital games research* 37 (2005), 2.
10. W. IJsselsteijn, K. Poels, and Y. de Kort. 2008. The Game Experience Questionnaire: Development of a self-report measure to assess player experiences of digital games. *TU Eindhoven, Eindhoven, The Netherlands* (2008).
11. C. Jennett, A. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton. 2008. Measuring and defining the experience of immersion in games. *International journal of human-computer studies* 66, 9 (2008), 641–661.
12. I. Nordin, A. Denisova, and P. Cairns. 2014. Too Many Questionnaires: Measuring Player Experience Whilst Playing Digital Games. *Seventh York Doctoral Symposium on Computer Science & Electronics* 69 (2014).
13. H. Qin, P.L.P. Rau, and G. Salvendy. 2009. Measuring player immersion in the computer game narrative. *Intl. Journal of Human–Computer Interaction* 25, 2 (2009), 107–133.
14. S. Rigby and R. Ryan. 2007. The player experience of need satisfaction (PENS) model. *Immersyve Inc* (2007).
15. R. Ryan, S. Rigby, and A. Przybylski. 2006. The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion* 30, 4 (2006), 344–360.
16. P. Sweetser and P. Wyeth. 2005. GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE)* 3, 3 (2005), 3–3.
17. B. Witmer and M. Singer. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and virtual environments* 7, 3 (1998), 225–240.