



Measuring engagement in video game-based environments: Investigation of the User Engagement Scale



Eric N. Wiebe^{a,*}, Allison Lamb^a, Megan Hardy^a, David Sharek^b

^a NC State University, Raleigh, NC, USA

^b Playgraph, Inc., USA

ARTICLE INFO

Article history:

Available online 27 December 2013

Keywords:

Engagement
Psychometric measurement
Flow
Video game

ABSTRACT

This research investigated the use of the User Engagement Scale (UES) as a psychometric tool to measure engagement during video game-play. Exploratory factor analysis revealed four factors (Focused Attention, Perceived Usability, Aesthetics, and Satisfaction) as compared to the six found in the original development of the UES. In the context of video game-play, a revised UES (UESz) demonstrated better psychometric properties than the original UES defined by six subscales, including enhanced reliability. Further validity analysis included comparisons with the Flow State Scale (FSS), showing the complementary nature of the two scales and what constructs both scales might be measuring in a video game context. Criterion validity analysis demonstrated that UESz was more predictive of game performance than the FSS. Findings related to both the UESz and FSS were discussed relative to an overarching framework of hedonic and utilitarian qualities of game-play.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

While cognitive theories and constructs continue to dominate much of the research in human–computer interaction, affective dimensions of the user experience have increasingly gained the attention of researchers (cf., Picard, 2010). In particular, the intersection of affective and cognitive dimensions are being studied as key underlying constructs to help explain the rapid rise of popularity of video games and their efficacy for generating extended task-oriented behaviour (Przybylski, Rigby, & Ryan, 2010). For many researchers, gaming is of particular interest because of its potential use as a guiding design heuristic in learning-based environments (e.g., Mayo, 2009; Whitton, 2011).

This increased interest in the intersection of affective and cognitive psychological dimensions of user experience with video games has carried over to other computational environments. At this nexus, engagement has been recognized as a key factor in understanding general user behaviour and overall efficacy of goal or task-oriented behaviour within computer-based environments; including work-oriented information retrieval tasks, social networking tools (e.g., Facebook™), games, traditional educational environments, and hybrid environments such as game-based learning (Boyle, Connolly, & Hainey, 2011; Faiola, Newlon, Pfaff,

& Smyslova, 2012; O'Brien and Toms, 2008). Evaluation and research on these types of computer-based environments have included measures of engagement, yet it is recognized that better tools are needed to help define and measure this construct (cf., Attfield, Kazai, Lalmas, & Piwowarski, 2011).

The goal of this study is to extend and refine existing work in defining user engagement as it relates to computer and game-based environments. This study focuses on extending the ongoing work being conducted by the team of O'Brien and Toms (2008, 2010, 2012) on developing a self-report instrument of user engagement. Their work is extended by investigating the use of their User Engagement Scale (UES) in the context of game-based environments. Additionally, this study continues the work on validating and refining the instrument and its underlying constructs.

1.1. Theoretical framework

In the past couple of decades, two related frameworks of engagement have been developed in the contexts of school (academic) settings and in human–computer interaction. School-based engagement is the broader of the two, encompassing an individuals' engagement with academic activities in school but influenced by factors both within and outside of school. Work by Appleton, Christenson, Kim, and Reschly (2006), Fredricks, Blumenfeld, and Paris (2004) have developed a sophisticated, multifaceted view of engagement that can be conceptualized at many different levels. Their complementary work demonstrates that family, community, culture and educational context are all important antecedent

* Corresponding author. Address: NC State University, Box 7801, Raleigh, NC 27695-7801, USA. Tel.: +1 919 515 1753.

E-mail addresses: eric_wiebe@ncsu.edu (E.N. Wiebe), alamb4@ncsu.edu (A. Lamb), rmhardy@ncsu.edu (M. Hardy), dsharek@playgraph.com (D. Sharek).

factors mediating engagement. Fredricks et al. (2004) describes factors related to engagement that are measurable at different contextual levels: school, classroom, and individual levels. A complementary approach by Appleton et al. (2006) identifies facets of engagement related to academic (time on task), behavioural (classroom participation), cognitive (strategizing), and psychological (belonging) settings. Both researchers note that engagement can happen in the context of group-social interaction and at the individual level. From this research comes an understanding that engagement can be effectively studied and measured at the individual level as individuals conduct cognitively demanding task-oriented activities. Performance is frequently considered as an outcome measure in cognitive tasks and is mediated by affective dimensions. Thus, engagement becomes an important factor to measure if performance is to be understood, because it helps explain the critical mediating role that affective dimensions play in cognitive tasks.

Of particular interest to this current study, Appleton, Fredricks and other allied researchers helped define where engagement resides with regards to a broad spectrum of cognitive and affective dimensions or states such as motivation and self-efficacy (cf., Sharek, 2012). They note that self-determination theory (Ryan & Deci, 2000) helps explain how intrinsic motivation—driven by factors such as the need for competence—is an important precursor for engagement (Boyle et al., 2011; Przybylski et al., 2010). However, while engagement is certainly related to motivation, it can be considered a separate construct (Appleton et al., 2006; Przybylski et al., 2010). One can conceive engagement as a series of (state-like) temporal interactions during task while motivation is a more (state or trait-like) global personal orientation towards the learning/task (Bempechat & Shernoff, 2012). The interaction between the two can form a feedback loop where the experience with the task can shape the more state-like elements of self-efficacy and motivation which in turn influences the user's desire to re-engage with a task (Sharek, 2012).

A second line of research has developed around understanding engagement in individual, task-oriented endeavors using computer-based tools (e.g., O'Brien and Toms, 2008). This work has developed a conceptual model of engagement in the context of human-computer interaction, primarily around tasks related to information search and retrieval, but also encompassing activities such as online shopping and video games (O'Brien and Toms, 2010). O'Brien and Toms have developed a model of engagement that is both a process and a product of interaction. It represents a cyclical experience of engagement and reflection on this interactive experience that helps shapes decisions about future engagement (O'Brien and Toms, 2010). While O'Brien and Toms (2008) provide a comprehensive review of the different theoretical sources of their identified facets of engagement, the work of Hassenzahl, Diefenbach, and Göritz (2010) can be used to provide a very succinct lens by organizing these experiences and motivations for engagement into two categories, or qualities, of the user experience. First, there are pragmatic qualities related to the usefulness and usability of the system. Second, there are the hedonic qualities of motivation, stimulation, and challenge for the user. The pragmatic qualities of utility/usefulness and usability have a long, established history in research in human-computer interaction (cf., Shneiderman, 1998) and form the backbone of well-established frameworks such as the Technology Acceptance Model (Bourgonjon, Valcke, Soetaert, & Schellens, 2010; Venkatesh & Bala, 2008). Positive responses by the user to both the usability and usefulness of a computer-based system for a task are seen as key prerequisites for user engagement (O'Brien and Toms, 2010).

Hassenzahl et al.'s (2010) hedonic qualities sit at the centre of what has become the areas of engagement that engender the most interest among researchers looking at video games and game-

based learning environments. It is the hedonic qualities of games that are seen as the key elements that explain the perception of high engagement by users (Malone & Lepper, 1987; Przybylski et al., 2010). One facet of this quality is the perceived aesthetics of the computer-based environment (Skelly, Fries, Linnett, Nass, & Reeves, 1994; Vorderer, Klimmt, & Ritterfeld, 2004). Aesthetics works at many different levels to both motivate initial interaction with a system, enjoyment while using it, and (perhaps) the perceived overall usability of the system. The notion of "play" encompasses many of the other facets of the hedonic quality of human-computer interaction, especially as they relate to game-based environments (Rieber, 1996; Stephenson, 1967). While play is often thought of as a physical activity, it can also represent interaction with virtual spaces. Also relevant is that while play can be task-oriented, these tasks can just as (or more) easily be driven by hedonic rather than pragmatic qualities of the experience. Though play in both physical and virtual spaces is typically thought of as an activity or behaviour, what is most relevant to engagement is perhaps the user's perceived experience while at play and their response to this experience.

Flow Theory (Csikszentmihalyi, 1990) is used by O'Brien and Toms (2008, 2010), along with numerous other researchers (e.g., Boyle et al., 2011; Sherry, 2004), to describe what an individual experiences during play and, thus, is a central theoretical frame for understanding user experience in game-based settings and as a means for explaining user engagement in such environments. Flow can be thought of as a deep immersive experience that results from an individual engaging in a task that has an appropriate balance of challenge relative to a user's skill level (Csikszentmihalyi, 1990; Faiola et al., 2012; Sherry, 2004). Flow and game play are often linked in contexts where the user finds a familiar formal structure (of the game) but novel content created by the system design and user choice within the system (Sherry, 2004). The connection here can be found within the larger framework of motivation and engagement where the user will leverage their familiarity and past experiences of the formal structure to determine whether the challenge being presented to them is both achievable and desirable. From this flow experience is a positive affective response—enjoyment and satisfaction—that leads an individual to both reflect positively on the experience and, typically, want to re-engage with it again (Sharek, 2012). Despite the connections that many have found between Flow Theory, as conceived by Csikszentmihalyi (1990), and game-based environments, researchers have noted that the connections between the two are often deceptively simple on the surface and decidedly complex in direct application in research (cf., Weibel, Wissmath, Habegger, Steiner, & Groner, 2008). Indeed, while flow and enjoyment reported in game play may be highly correlated, they might be considered separate constructs to be measured and characterised individually (Boyle et al., 2011; Weibel et al., 2008).

1.2. Measurement

This emerging framework for understanding engagement leads to the use of multiple measurement paradigms and approaches to measuring data arising from users in task-oriented contexts. Given that the larger framework of engagement encompasses both behaviours in physical and virtual spaces and the resulting psychological states, it is appropriate that researchers have employed varied and often multiple measurement approaches. On the behavioural side, direct observational measures have been used to characterise overt behaviour in educational and other settings (Admiraal, Huijenga, Akkerman, & Dam, 2011). In the context of computer-based activities, in addition to direct observation of activity, trace/log data from interaction from the system is also employed (Lehmann, Lalmas, Yom-Tov, & Dupret, 2012). To

characterise the psychological state of users, physiological measurement has become increasingly popular due to the rapid technological enhancement of measurement and analysis tools along with theoretical models to guide interpretation (Hardy, Wiebe, Grafsgaard, Boyer, & Lester, 2013; Mandryk, Inkpen, & Calvert, 2006).

However, self-report measures (usually post hoc) continue to be the most popular measure for characterizing the psychological state of user engagement. Previously cited works on the development of a framework of user engagement note the importance of using multiple measures of engagement to capture behavioural and psychological aspects of engagement, but also to capture the differing facets of psychological engagement (Appleton et al., 2006; Boyle et al., 2011). Self-report measures continue to represent a robust, easy to implement approach to collecting valid, reliable data for research and evaluation of user experiences in video game-based environments. Despite the general interest in engagement, there is a paucity of well-validated self-report instruments for measuring this construct. Appleton et al. (2006) developed the Student Engagement Instrument—closely related to both his theoretical frame and that of Fredricks et al. (2004)—however, this was developed for use in the context of schools and school-based activities. There have been other instruments developed especially for game-based environments, often employing flow and other engagement-related constructs, but they have not been widely used or validated (e.g., Brockmyer et al., 2009; Fu, Su, & Yu, 2009). More widely used in game-based research has been an instrument developed around Csikszentmihalyi's (1990) original conception of flow for use in the context of physical activity, the Flow State Scale (Jackson & Marsh, 1996; Marsh & Jackson, 1999). Interestingly, while this scale was developed and validated for measurement of sports-related physical activity, it has been used to measure a broad conceptual definition of flow in computer-based activities, including video games (e.g., Kato & Suzuki, 2003; Liao, 2006; Ma & Williams, 2011). The appeal of using the Flow State Scale in contexts such as game-based environments may come both from the broad assessment of the flow experience (the scale has nine separate subscales) and the general wording of the scale items (i.e., not specifically tied to sports activities) (Jackson & Marsh, 1996).

Perhaps most promising in this area of self-report of game-based engagement has been the development and refinement of the User Engagement Scale by O'Brien and Toms (2008, 2010, 2012). As previously noted, O'Brien and Toms used work in educational multimedia as a starting point to define a number of both pragmatic and hedonic dimensions that influence user engagement. In addition to the construct of flow, they also wanted to capture the hedonic aspects of aesthetics, fun/pleasure, and novelty along with pragmatic aspects of usability and a more reflective sense of whether the user would like to re-engage with this experience again in the future. From these constructs, they developed and validated a self-report instrument with six subscales representing six unique constructs (O'Brien and Toms, 2010):

- AE – Aesthetics
 - Visual appearance
- EN- Endurability
 - Holistic response to experience – reflective-forward: would I use it again and recommend to others?
- FI - Felt Involvement
 - Interesting and fun experience?
- FA - Focused Attention
 - Based on Flow Theory (focused concentration, absorption, temporal dissociation)
- NO – Novelty
 - Interest and curiosity evoked by the software system

- PU - Perceived Usability
 - Both affective (frustration) and cognitive (effortful) aspects of use of the system

As part of their original validation work, they interviewed users about their experiences in four different contexts, video games, on-line learning, Internet web searching, and shopping. They then conducted both an exploratory and confirmatory factor analysis with the instrument in the context of an online shopping environment (O'Brien and Toms, 2010). Other researchers have since used the instrument in the context of interactive information retrieval (Arguello, Wu, Kelly, & Edwards, 2012), online multiplayer games and avatar creation (Turkay, 2012), and social networking (i.e., Facebook) (Banhawi, Ali, & Judi, 2012).

More recently, O'Brien and Toms have revisited their instrument in light of research done by others with the instrument, and their own further investigations with it in the context of information retrieval (O'Brien and Toms, 2012). One of the most interesting findings reported in this study is the question of whether Endurability (EN), Novelty (NO), and Felt Involvement (FI) exist as three separate, or just one single factor. While the original validation in the context of online shopping confirmed these six factors, a further study on online search (O'Brien, 2010) as well as a social networking study (Banhawi et al., 2012) revealed factor loadings that had these three merged as one. As O'Brien (2010) notes, the context of the task can have strong influences on one's motivations and thus how they perceive their engagement. Just as important, they note, this is still a relatively new scale in need of further research in multiple contexts. O'Brien and Toms (2012) further work with online information retrieval has confirmed the stability of the Aesthetics (AE), Focused Attention (FA), and Perceived Usability (PU) scales, but has continued to document the relative fluidity of items from the other three subscales, based on the context of the task.

Published work on the UES points to the promise of the scale for use in investigating engagement in computer-based environments, including game-based ones. Flow Theory, which has been used by many researchers as a lens to examine the experience of game-play is explicitly measured via the subscale of Focused Attention, along with other relevant hedonic and pragmatic dimensions of experience. This approach contrasts to the development of the Flow State Scale, which focused exclusively on the measurement of the construct of flow. However, the UES has not been specifically validated in the context of game-play. In addition, prior work has pointed to questions as to exactly how many unique constructs are being measured when used in other computer-based contexts. This current work will specifically address the question of, in the context of game-play, how many distinct factors emerge through an exploratory factor analysis. In addition, ability of the different UES factors to predict game play performance will be explored and compared to the Flow State Scale.

2. Results

2.1. Participants

A total of 572 participants were recruited via Mechanical Turk (Amazon.com., 2010). Prior research has indicated that Mechanical Turk is a reliable source of subject participants (Behrend, Sharek, Meade, & Wiebe, 2011). Data for 159 participants was eliminated due to incomplete data, nonvariance of response, failure to meet minimum playtime or indication that they were unable to complete the game portion of the study. The remaining 413 participants ranged in age from 18 to 66 ($M=29$, $SD=9.75$). The education level reached by participants ranged from high school

degree (8%), college degree (31%) to graduate degree (36%) with the remaining reporting some college or graduate schooling (24%). A subset of the participants ($N = 186$) completed both the UES and Flow State Scale (FSS). This subset had nearly identical demographic characteristics to the sample as a whole.

2.2. Stimuli and measures

Participants played an online strategy game called Block Walk (Sharek, 2009) and were then linked to the survey portion of the study. The goal of the tile-based game (see Fig. 1) is to navigate a rectangular block around obstacles, so that it ends up placed vertically on top of a goal. Block Walk was chosen because prior studies had found it to be an engaging task (using other metrics of engagement) and because it requires strategic thinking and spatial evaluation that is relevant in educational research (Sharek, 2012).

2.2.1. User Engagement Scale

The User Engagement Scale (UES) was measured as a 5-point Likert scale and is comprised of 31 items and six subscales: focused attention (FA), felt involvement (FI), novelty (NO), endurance (EN), aesthetics (AE), and perceived usability (PU). These items were modified for use in a gaming context (example: original FA5 “The time I spent shopping just slipped away”, modified “The time I spent playing the game just slipped away”). The reliability for the scale as a whole was acceptable (Cronbach's $\alpha = .92$). The sample population ($N = 413$) provided a ratio of 13 cases per variable.

2.2.2. Flow State Scale

The Flow State Scale was measured as a 5-point Likert scale and is comprised of 36 items and nine subscales: challenge-skill balance, action-awareness merging, clear goals, unambiguous feedback, concentration on task at hand, paradox of control, loss of self-consciousness, transformation of time, and autotelic experience. The reliability of the scale was acceptable (Cronbach's $\alpha = .93$).

2.2.3. Procedure

Participants were required to play the game for 10 min, but were given the option to continue playing past the required

amount of time (Min = 10, Max = 90; $M = 20$). After the minimum time limit was reached, they could choose to exit the game and begin the survey. The survey consisted of demographic questions, a question regarding whether the game functioned sufficiently, the UES, the FSS (for a subsample of participants), and some usability questions created specifically for the study.

2.3. Exploratory factor analysis

Multiple criteria were used to determine the factorability of the 31 items in the UES. First, the sample size was considered. Kass and Tinsley (1979) recommend that 300 or greater participants tend to be stable regardless of participant to variable ratio. Others agree that an N of 300 or more as being suitable for factor analysis (Comrey & Lee, 1992; Tabachnick & Fidell, 2007). Thus, our sample size ($N = 413$) was deemed more than adequate.

Second, the correlation matrix between the items was examined. Tabachnick and Fidell (2007) recommend that all correlation coefficients be above .3. Similarly, Field (2009) recommends eliminating variables that do not correlate highly with any other variables. In our correlation matrix, the item “I felt in control of my gaming experience” (PU7) – did not correlate above .3 with any other item. Therefore, it was removed from subsequent analyses. Third, the Kaiser–Meyer–Olkin (KMO) is a measure of sampling adequacy, with values above .9 being considered superb (Hutcheson & Sofroniou, 1999). Our KMO was .926, a value indicating that the patterns of correlations were compact and a factor analysis should yield distinct and reliable factors (Field, 2009). Additionally, Bartlett's test of sphericity was significant ($\chi^2(435) = 6811.01$; $p < .001$), and the diagonals of the anti-correlation matrix were all above .8. Finally, communalities are mostly within the .5 range. MacCallum, Wildaman, Zhang, and Hong (1999) indicate that with a sample size greater than 200, communalities around .5 are sufficient. Overall, based on the consideration of these criteria, factor analysis was determined to be suitable.

2.3.1. Extraction method

Principle Axis Factoring (PAF) with a Promax rotation was used to estimate factors underlying the construct of engagement. Eigen values show that five factors explained 32.62%, 14.53%, 7.08%, 4.20%, and 3.53% of the variance, respectively (see Table 1). After



Fig. 1. Screenshot of game play in the video game, Block Walk (Sharek, 2009).

Table 1
Percent variance explained from factors with eigenvalues over 1.

Factor	Initial eigenvalues		
	Total	% Of variance	Cumulative (%)
1	9.785	32.618	32.618
2	4.358	14.528	47.146
3	2.124	7.080	54.226
4	1.261	4.204	58.430
5	1.060	3.532	61.962

rotation, one item (EN2)—“I consider my gaming experience a success”—was the only item loading on the fifth factor. Thus, this item and respective factor were eliminated. Additionally, the item (FI2)—“I felt involved in this gaming task.”—was eliminated because it did not contribute to the factor structure and had factor loadings below .4. Eigen values and cumulative variance explained can be found in Table 1. A scree plot can be found in Fig. 2. Factor loadings were generally consistent with results found by O'Brien and Toms (2012) and are compared to the original UES subscales in Table 2.

Factor 1 Focused Attention (FAz). This factor accounted for 33% of the variance and consisted of the items (FA1-7, and FI1). This factor was the original UES's Focused Attention subscale with the addition of one Felt Involvement item.

Factor 2 Perceived Usability (PUz). This factor accounted for 15% of the variance and consisted of the items (PU1-6, PU8, and EN3). This factor was essentially the original UES's Perceived Usability subscale with the addition of one Endurability item.

Factor 3 Aesthetics(AEz). This factor accounted for 7% of the variance and consisted of the items (AE1-5). This factor was the original Aesthetics subscale.

Factor 4 Satisfaction (SAz). This factor accounted for 4% of the variance and consisted of the items (NO1-3, EN1, EN4, EN5, and FI3). This factor was a combination of items from the original UES's subscales of Novelty, Endurability, and one Felt Involvement item.

2.4. Reliability analysis for subscales

The revised User Engagement Scale emerging from the factors found through EFA (UESz) was measured as a 5-point Likert scale and is comprised of 28 items and four subscales: focused attention (FAz), perceived usability (PUz), aesthetics (AEz), and satisfaction (SAz). The reliability for the modified scale as a whole was acceptable (Cronbach's alpha = .91).

Reliability analyses and correlations were performed both for the original UES subscales (see Tables 3 and 4) and for the factors as determined by our EFA (UESz) (see Tables 5 and 6) as a means of comparing the associations seen in the original UES subscales and the four factor model found in the EFA (UESz). Means were calculated by summing participants' ratings on a 5-point scale of items within each subscale and dividing by the total number of items for that subscale. These individual scores were then calculated to obtain means and standard deviations for each subscale.

Note that the Cronbach's alpha is very close or identical between the original FA, PU, and AE subscales and the UESz Factors 1 through 3 (FAz, PUz, and AEz) of the EFA. Factor 4 (SAz) has an improved Cronbach's alpha (.88) over the original FI (.77), NO (.81), and EN (.67) subscales. The UESz factors show a similar correlation pattern to the original subscales with Perceived Usability showing the lowest (though still significant) correlations with Focused Attention and Aesthetics. Factor 4 (SAz) shows the highest correlations with FAz and AEz; this parallels the relatively high correlations in the original UES of FI, NO, and EN with FA and AE. FI, NO, and EN also had relatively high correlations with each other in the original UES.

2.5. Analysis of UESz with FSS

Similar to the comparison of the UESz to the original UES, associations of the four factors derived as part of the EFA (UESz) were compared directly to the Flow State Scale (Jackson & Marsh, 1996) in Table 7. In this inter-correlation table, the relatively low correlations between PUz and both FAz and AEz seen in Table 6 follow the same trend and become non-significant in Table 7. The FSS

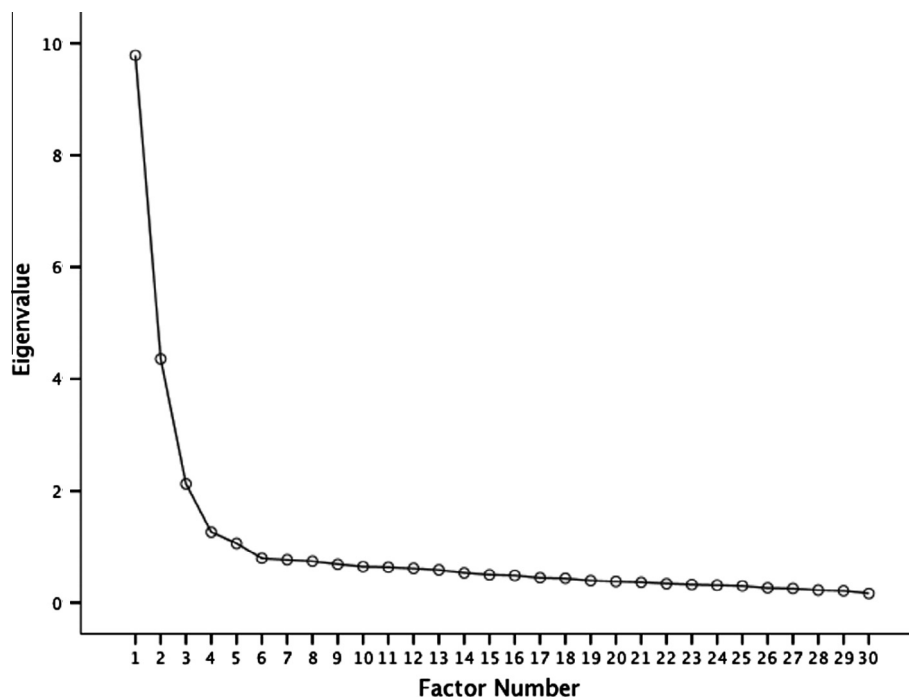


Fig. 2. Scree plot of eigenvalues for each factor.

Table 2
Principle axis factoring with promax rotation.

	UES Sub- scale	*	Factor			
			1	2	3	4
When I was playing the game, I lost track of the world around me.	FA4		.881	-.005	-.007	.040
I blocked out things around me when I was playing the game on this website.	FA3		.777	.018	-.021	-.036
The time I spent playing the game just slipped away.	FA5		.727	-.013	-.018	-.014
I was absorbed in my gaming task.	FA6		.638	.027	.180	-.140
I was so involved in my gaming task that I lost track of time.	FA2		.638	.008	-.126	.099
During this gaming experience I let myself go.	FA7		.605	-.046	.045	.076
I lost myself in this gaming experience.	FA1		.501	.003	-.018	.126
I was really drawn into my gaming task.	FI1		.484	-.018	-.032	.061
I felt discouraged while on the website.	PU4		.024	.879	.002	.087
I felt annoyed while visiting the website.	PU3		-.052	.824	.126	.041
Using the website was mentally taxing.	PU5		.098	.816	-.065	.016
I found the website confusing to use.	PU2		.060	.785	.067	-.086
I felt frustrated while visiting the website.	PU1		-.031	.703	.002	.025
I could not do some of the things I needed to do on the gaming website.	PU8		-.020	.563	-.162	-.095
The gaming experience was demanding.	PU6		-.069	.551	.028	.028
This gaming experience did not work out the way I had planned.	EN3		-.048	.437	-.118	-.140
I liked the graphics and images used on the website.	AE3		-.036	.054	.815	.038
The website appealed to my visual senses.	AE4		.027	-.028	.810	-.040
The website was aesthetically appealing.	AE2		-.081	-.017	.790	.090
The screen layout of the website was visually pleasing.	AE5		.146	-.064	.710	-.144
The website was attractive.	AE1		-.063	-.018	.703	.130
The content of the gaming website incited my curiosity.	NO2		.059	-.059	-.036	.825
I would continue to go to this website out of curiosity.	NO1		.054	-.055	-.011	.807
I would recommend playing the game on the website to my friends and family.	EN5		.005	.000	.071	.679
Playing the game on this website was worthwhile.	EN1		.136	.004	.090	.502
I felt interested in my gaming task.	NO3		.026	.093	.090	.478
My gaming experience was rewarding.	EN4		.019	-.041	-.045	.457
This gaming experience was fun.	FI3		.042	.121	.048	.441

Note. green = expected, orange = not expected, white = item not included.

*results found from EFA in O'Brien & Toms (2012).

is significantly correlated with all four factors of the UESz, though to a smaller extent with PUz.

2.6. Regression analyses to predict performance

2.6.1. Multiple regression with UESz subscales

Criterion validity was explored by examining the degree to which either the UESz subscales or the FSS were able to predict performance in the game. It is assumed that individuals with higher self-reported engagement should perform better at the game (Przybylski et al., 2010). The maximum level of difficulty reached

by the participant in the game was used as the best proxy for performance (Sharek, 2012). Since the FSS and UESz were developed with overlapping but not identical conceptual models (borne out in the intercorrelation table, Table 7), these two scales were also used together to examine their collective predictive power.

Initially, a multiple regression was conducted to predict the maximum level of difficulty reached (Max Level) from the scores of the newly derived subscales of UESz (Table 8). The assumptions of linearity, independence of errors, and homoscedasticity were met. Additionally, there was independence of residuals, as assessed by a Durbin-Watson statistic of 2.07. Results show that this model

Table 3
Reliability analysis for subscales from original UES.

Subscale	Number Items	Cronbach's alpha	Mean	Standard Deviation
Focused Attention (FA)	7	.87	3.52	0.73
Perceived Usability (PU) [*]	7	.85	3.55	0.78
Aesthetics (AE)	5	.88	3.60	0.78
Felt Involvement (FI)	3	.77	3.89	0.71
Novelty (NO)	3	.81	3.67	0.78
Endurability (EN) ^{**}	4	.67	3.54	0.63

^{*} If PU7 is removed, Cronbach's alpha increases to .89.
^{**} If EN3 is removed, Cronbach's alpha increases to .77.

Table 4
Correlations between subscales from original UES.

Subscale	FA	PU	AE	FI	NO	EN
Focused Attention (FA)	–					
Perceived Usability (PU)	.12 [*]	–				
Aesthetics (AE)	.46 ^{***}	.14 ^{**}	–			
Felt Involvement (FI)	.69 ^{***}	.27 ^{***}	.52 ^{***}	–		
Novelty (NO)	.59 ^{***}	.17 ^{***}	.59 ^{***}	.67 ^{***}	–	
Endurability (EN)	.55 ^{***}	.35 ^{***}	.52 ^{***}	.66 ^{***}	.66 ^{***}	–

Note. N = 413.
^{*} p < .05.
^{**} p < .01.
^{***} p < .001.

Table 5
Reliability analysis for subscales determined by EFA (UESz).

Subscale	Number items	Cronbach's alpha	Mean	Standard deviation
Factor 1. Focused Attention (FAz)	8	.88	3.55	0.72
Factor 2. Perceived Usability (PUz)	8	.89	3.51	0.84
Factor 3. Aesthetics (AEz)	5	.88	3.60	0.78
Factor 4. Satisfaction (SAz)	7	.88	3.70	0.70

Table 6
Correlations between subscales determined by EFA (UESz).

Subscale	Factor			
	1	2	3	4
Factor 1. Focused Attention (FAz)	–			
Factor 2. Perceived Usability (PUz)	.105 [*]	–		
Factor 3. Aesthetics (AEz)	.472 ^{***}	.104 [*]	–	
Factor 4. Satisfaction (SAz)	.652 ^{***}	.183 ^{***}	.630 ^{***}	–

Note. N = 413.
^{**} p < .01.
^{*} p < .05.
^{***} p < .001.

significantly predicted Max Level $F(4, 181) = 6.585, p < .001$. Overall, this model accounted for 11% of the variance in Max Level. FAz, with a β of .078 was not significant in this model. All other subscales were significant. Also of note, AEz has a negative β .

2.6.2. Linear regression with FSS

Next, a linear regression was run to predict Max Level from scores on the FSS (Table 9). With this model, there was also independence of residuals, as assessed by a Durbin-Watson statistic of 2.14. Results show that the FSS significantly predict max level

Table 7
Correlations between subscales from UESz and FSS.

Subscale	FAz	PUz	AEz	SAz	FSS
Focused Attention (FAz)	–				
Perceived Usability (PUz)	.13	–			
Aesthetics (AEz)	.47 ^{***}	.06	–		
Satisfaction (SAz)	.63 ^{***}	.18 ^{**}	.65 ^{***}	–	
Flow State Scale (FSS)	.61 ^{***}	.18 ^{**}	.58 ^{***}	.75 ^{***}	–

Note. N = 186.
^{**} p < .01.
^{***} p < .001.

Table 8
Multiple regression predicting max level from UESz subscales.

Subscale	B (SE _B)	β
Focused Attention (FAz)	.568 (.613)	.078
Perceived Usability (PUz)	1.043 (.452)	.164 [*]
Aesthetics (AEz)	–2.125 (.577)	–.340 ^{***}
Satisfaction (SAz)	2.092 (.729)	.305 ^{**}

Note. N = 186.
^{*} p < .05.
^{**} p < .01.
^{***} p < .001.

Table 9
Linear regression predicting max level from FSS.

	B (SE _B)	β
FSS	2.187 (.742)	.212 ^{**}

Note. N = 186.
^{*} p < .05.
^{***} p < .001.
^{**} p < .01.

$F(1, 184) = 8.690, p = .004$. Overall, scores on the FSS accounted for 4% of the variance in max level reached.

2.6.3. Hierarchical multiple regression with UESz subscales and FSS

Finally, a hierarchical multiple regression was run to determine if the addition of FSS improved the prediction of Max Level over and above the UESz subscales (Table 10). For this model, there was independence of residuals, as assessed by a Durbin-Watson statistic of 2.15. Tolerance values were greater than .1. Respectively, VIF values were less than 10. Model 1 contains just the UESz subscales while Model 2 contains both the UESz subscales and the FSS. The full model of UESz subscales and FSS (Model 2) was statistically significant $F(5, 180) = 6.509, p < .001, \text{adjusted } R^2 = .13$. The addition of FSS led to a statistically significant increase in R^2 of .026. That is, the addition of FSS explained 2.6% more of the variance in Max Level than the UESz subscales alone and that this R^2 change of .026 from Model 1 to Model 2 was significant, $F(1, 180) = 5.541, p < .05$. FSS was a significant contributor to Model 2 with a β of .257. FAz continued to be non-significant in Model 2, with its β dropping from .078 to .024. SAz moved from significant to non-significant between Model 1 and Model 2, with its β dropping from .305 to .172. PUz and AEz were significant in both models with only small shifts in β between the models.

3. Discussion

This study investigated the User Engagement Scale (UES), both in its original form as constructed by O'Brien and Toms (2010) and as a revised instrument based on a new exploratory factor analysis (UESz) conducted in the context of video game play. In

Table 10
Hierarchical regression predicting max level.

	Max level			
	Model 1		Model 2	
	<i>B</i> (<i>SE_B</i>)	β	<i>B</i> (<i>SE_B</i>)	β
Constant	7.016 (2.353)**		3.204 (2.833)	
Focused Attention (FAz)	.568 (.613)	.078	.160 (.626)	.024
Perceived Usability (PUz)	1.043 (.452)*	.164	.956 (.448)*	.150
Aesthetics (AEz)	−2.125 (.577)**	−.340	−2.345 (.577)**	−.375
Satisfaction (SAz)	2.092 (.729)**	.305	1.179 (.817)	.172
FSS			2.649 (1.125)*	.257
<i>R</i> ²	.108		.130	
<i>F</i>	6.585***		6.509***	
ΔR^2			.026	
ΔF			5.541*	

Note. *N* = 186.

* *p* < .05.

** *p* < .01.

*** *p* < .001.

addition, the UESz was compared to another instrument, the Flow State Scale, currently in use to measure engagement during video game play.

3.1. Factor structure

In the exploratory factor analysis, four factors emerged, three of which were well aligned with subscales previously constructed by O'Brien and Toms (2010) as part of the UES: Focused Attention (FAz), Perceived Usability (PUz), and Aesthetics (AEz). The fourth factor, called Satisfaction (SAz), was a combination of items from the original Endurability, Novelty, and Felt Involvement UES subscales. In addition, three items from the original scale were removed due to lack of correlation with other items, and factor loadings below .4 on these four factors. The newly emergent factor, called Satisfaction (SAz), corresponds to findings reported in O'Brien and Toms (2012) on work conducted by the these authors and others in the contexts of online shopping, online search, and Facebook use (Banhawi et al., 2012; O'Brien, 2010). This current study provides further evidence that perhaps the original constructs of Endurability, Novelty, and Felt Involvement should be considered a single construct in the context of these types of tasks.

The name Satisfaction was given to this new factor since a review of the items that load against this scale all relate closely to self-reported reflection on the more hedonic aspects of their experience. They were items that related to how fun and novel the experience was, and the likelihood that they would want to play/use this game again. This contrasted with the Perceived Usability (PUz) scale which clearly targeted the utilitarian aspects of the game environment. Aesthetics, focusing on the visual elements of the interface, is seemingly hedonic in nature. However, as Lavie and Tractinsky (2004) point out, an aesthetic response can be influenced by both the more classic elements of design—which are closely aligned with usability conventions—and the purely affective response to the visual design – even if the visual design runs counter to good usability principles. Possibly for this reason, Aesthetics emerges as a dimension that is related yet distinct from the set of attributes leading to satisfaction with use. Focused Attention (FAz) was also a distinct factor which had items that attempted to capture self-report of what the user was feeling while playing the game. FAz items such as “I was absorbed in my gaming task” seem quite different in temporal point of view from the more reflective items in the SAz factor: “My gaming experience was rewarding.” Focused Attention was similarly distinct from Aesthetics since a visual assessment of the interface could be made without any interaction with the game. While user assessment of Perceived Usability

also depends on interaction with the game, this construct falls on the utilitarian side of the utilitarian/hedonic continuum.

3.2. Reliability analysis

Reliability analysis of both the original UES and the newly derived UESz provides additional evidence that perhaps the original subscales of FI, NO, and EN should both shed a few items and be combined into a new subscale, SAz. The original FI and NO subscales only contained three items and EN four items; item counts considered small for a stable subscale (Clark & Watson, 1995). These three subscales also had the weakest Cronbach's alpha values. The new factor structure removed two items, sent one FI item to FAz and one EN item to PUz, and combined the remaining 7 items into the new SAz subscale. This not only provided an appropriate number of items in each of the new subscales (Furr, 2011), but also resulted in Cronbach's alpha values between .88 and .89 for all four subscales. The new subscales were also all significantly correlated with each other, but none over .65, meeting a target range of correlations that is typically considered reasonable. These correlations also reinforce the hedonic/utilitarian divide in the subscales, with PUz showing the lowest correlations with the other subscales.

3.3. Validity analysis of UESz

The inter-correlation matrix in Table 7 for the UESz and FSS provides some modest convergent and divergent validity for the UESz scale. The FSS, as conceptualized by its creators (Jackson & Marsh, 1996; Marsh & Jackson, 1999), is meant to be a holistic measure of a user's flow state. As such, it encompasses both the more narrow definition of flow as represented in the Focused Attention (FAz) subscale, but also the other hedonic dimensions of the UESz: Aesthetics (AEz) and Satisfaction (SAz). Similarly, the FSS would not be expected to capture the utilitarian dimension of usability (PUz). FSS-PUz correlation was still significant but quite a bit lower than with the other three UESz subscales. Interestingly, the inter-correlation model which included the FSS made the relatively modest correlations within the UESz between FAz-PUz and AEz-PUz non-significant; thus continuing to reinforce the hedonic/utilitarian distinctions within the UESz.

Finally, criterion validity of the UESz and FSS was explored by looking at each of these instruments' predictive power in a game-based context. As noted, it is assumed that individuals with higher self-reported engagement should perform better at the game (Przybylski et al., 2010), as defined as the maximum level

of difficulty reached by the players. It was assumed that both total amount of time played and the efficiency that players were able to move through levels would be key factors in the maximum game level reached. The regression analysis found that when all of the UESz subscales and the FSS were used together in a model and analysed holistically, they were significant predictors of Max Level. In addition, the UESz subscales predicted more variance (11%) than the FSS (4%). Coming out of these two regressions were a pair of unexpected findings. First, while the regression models were significant, they accounted for a somewhat disappointingly low amount of variance. Clearly, there are many other factors other than engagement that was determining this performance metric. One possibility is that Max Level is a less than ideal measure of performance. Also surprising was that FAz was not significant in the UESz-only regression model. It could very well be that while players were able to adequately report on their state of flow, entering a flow state was not necessary to perform well on this game. Finally, it was interesting to note that AEz had a negative Beta in the regression models. It could be that high ratings of aesthetics were actually a distraction in effective game-play. This interpretation would be supported by multimedia research related to extraneous cognitive load (cf., Mayer, 2003), and would also support a model of aesthetic response more based around non-utilitarian dimensions.

The hierarchical regression model in Table 10 provided further support for the interpretation of the relationship of the UESz subscales to each other, and to the FSS based on the inter-correlation matrices (Tables 6 and 7). Adding the FSS to the regression model offers a modest but significant improvement of the predictive power of the model (now at 13% of variance explained). Perhaps more important is that in this combined model, the FSS plus the PUz and AEz subscales are the significant contributors to the model. FAz, which was not significant in the original regression, is still not significant here in the combined model. In addition SAz, which was significant in the UESz-only model, is no longer significant in the combined model. The adjusted Beta drops by two-thirds for FAz, and almost in half for SAz between the two models. This reinforces the notion that the FSS is capturing elements of the FAz and SAz subscales, but provides a dimension unique from PUz and AEz when it comes to predicting performance in this game. It is worth noting that while the more narrowly defined construct of flow in the FAz subscale was not significant in this model, the more broadly defined FSS was. It leads to the conclusion that its contribution in this model is more important with regards to this notion of satisfaction with game play.

4. Conclusions

This study has contributed to a further understanding of the measurement of engagement in video game-based environments. Most importantly, it contributes the refinement of the original UES instrument (O'Brien and Toms, 2010) and suggests a modified version that is better optimized for use in video game research. Finally, it provides further scholarship on the constructs of engagement as they relate to video game-play, in particular the theoretical conception of flow and how it is experienced in such a context. In particular, when placed within Hassenzähl et al.'s (2010) framework of hedonic and utilitarian qualities, flow, satisfaction, usability, and aesthetics all provide unique contributions to the self-report of engagement during video game-play, with the former two considered hedonic and the latter two utilitarian. In addition, using this definition of these four constructs, the FSS (Jackson & Marsh, 1996) is measuring not just a more narrow interpretation of flow, but a broader hedonic quality that also includes satisfaction. While different

names can be attached to these constructs, the important finding is that unique qualities related to engagement are found along all four dimensions. On the other hand, in the context of video game-play, less support was found for attempting to measure along the six dimensions contained in the original UES. It is important to note that other studies measuring the psychometric characteristics of the UES have also suggested fewer dimensions.

While the validity analysis provided evidence of the general strength of the modified UES (UESz)—both in its own right and relative to the FSS—there were a number of questions raised as to what the FAz scale was or was not measuring and whether one should be expecting individuals to retrospectively report entering a state of flow in video games such as the one used in this study. Clearly more work needs to be done concerning what is appropriate criterion measures for flow, in particular, and engagement in video game-play more generally. Certainly more research is also needed in validating the UESz in more contexts; both video game play and other computer-based environments where the measurement of engagement is of interest to the research community. This work will help build a richer conceptualization of both flow and engagement more generally, frame these constructs relative to related constructs such as immersion (cf., Jennett et al., 2008), and add to a framework for evaluating the utility of other measurement techniques in video game-play research. Finally, another important line of work would be to use the measurement of engagement with the UESz as a means of exploring individual differences. This study did not collect deep enough data on the participants to effectively study such questions, but future work can pair the UESz with deep demographic profiles to explore how interface characteristics of video games interact with individual differences to create differing levels of engagement.

Acknowledgements

The authors would like to acknowledge that Jennifer London provided help on the initial data collection and analysis of this research project. This material is based upon work partially supported by the National Science Foundation under Grant No. DRL-1007962.

References

- Admiraal, W., Huizenga, J., Akkerman, S., & Dam, G. T. (2011). The concept of flow in collaborative game-based learning. *Computers in Human Behavior*, 27(3), 1185–1194. <http://dx.doi.org/10.1016/j.chb.2010.12.013>.
- Amazon.com. (2010). Mechanical Turk. <<http://www.mturk.com/>> (retrieved 21.07.011).
- Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the student engagement instrument. *Journal of School Psychology*, 44(5), 427–445.
- Arguello, J., Wu, W.-C., Kelly, D., & Edwards, A. (2012). Task complexity, vertical display and user interaction in aggregated search. In *Paper presented at the proceedings of the 35th international ACM SIGIR conference on research and development in information retrieval*, Portland, Oregon, USA.
- Attfeld, S., Kazai, G., Lalmas, M., & Piwowarski, B. (2011). Towards a science of user engagement (Position paper). In *Paper presented at the WSDM workshop on user modelling for web applications*, Hong Kong, China.
- Banhawi, F., Ali, N. M., & Judi, H. M. (2012). User engagement attributes and levels in facebook. *Journal of Theoretical and Applied Information Technology*, 41(1), 11–19.
- Behrend, T. S., Sharek, D. S., Meade, A. W., & Wiebe, E. N. (2011). The viability of crowdsourcing for survey research. *Behavior Research Methods*. <http://dx.doi.org/10.3758/s13428-011-0081-0>.
- Bempechat, J., & Shernoff, D. (2012). Parental influences on achievement motivation and student engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 315–342). US: Springer.
- Bourgonjon, J., Valcke, M., Soetaert, R., & Schellens, T. (2010). Students' perceptions about the use of video games in the classroom. *Computers and Education*, 54(4), 1145–1156. <http://dx.doi.org/10.1016/j.compedu.2009.10.022>.

- Boyle, E., Connolly, T. M., & Hainey, T. (2011). The role of psychology in understanding the impact of computer games. *Entertainment Computing*, 2(2), 69–74. <http://dx.doi.org/10.1016/j.entcom.2010.12.002>.
- Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., & Pidruzny, J. N. (2009). The development of the game engagement questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology*, 45(4), 624–634. <http://dx.doi.org/10.1016/j.jesp.2009.02.016>.
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309–319.
- Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Hillsdale, N.J.: Erlbaum.
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York, NY: Harper-Row.
- Faiola, A., Newlon, C., Pfaff, M., & Smyslova, O. (2012). Correlating the effects of flow and telepresence in virtual worlds: Enhancing our understanding of user behavior in game-based learning. *Computers in Human Behavior*, 29(3), 1113–1121. <http://dx.doi.org/10.1016/j.chb.2012.10.003>.
- Field, A. P. (2009). *Discovering statistics using SPSS* (3rd ed.). Los Angeles: Sage.
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59–109.
- Fu, F.-L., Su, R.-C., & Yu, S.-C. (2009). EGameFlow: A scale to measure learners' enjoyment of e-learning games. *Computers and Education*, 52, 101–112. <http://dx.doi.org/10.1016/j.compedu.2008.07.004>.
- Furr, M. R. (2011). *Scale construction and psychometrics for social and personality psychology*. London, UK: Sage.
- Hardy, M., Wiebe, E. N., Grafsgaard, J. F., Boyer, K. E., Lester, J. C. (2013). Physiological responses to events during training: Use of skin conductance to design adaptive learning systems. In *Proceedings of the human factors and ergonomic society 57th annual meeting*.
- Hassenzahl, M., Diefenbach, S., & Göritz, A. (2010). Needs, affect, and interactive products – Facets of user experience. *Interacting With Computers*, 22(5), 353–362.
- Hutcheson, G. D., & Sofroniou, N. (1999). *The multivariate social scientist*. London: Sage.
- Jackson, S. A., & Marsh, H. W. (1996). Development and validation of a scale to measure optimal experience: The flow state scale. *Journal of Sport and Exercise Psychology*, 18, 17–35.
- Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., et al. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, 66(9), 641–661.
- Kass, R. A., & Tinsley, H. E. A. (1979). Factor analysis. *Journal of Leisure Research*, 11, 120–138.
- Kato, Y., & Suzuki, K. (2003). An approach for redesigning learning environments with Flow Theory. In D. Lassner & C. McNaught (Eds.), *Proceedings of world conference on educational multimedia, hypermedia and telecommunications* (pp. 281–288).
- Lavie, T., & Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of web sites. *International Journal of Human-Computer Studies*, 60(3), 269–298.
- Lehmann, J., Lalmas, M., Yom-Tov, E., & Dupret, G. (2012). Models of user engagement. In *Paper presented at the proceedings of the 20th international conference on user modeling, adaptation, and personalization, Montreal, Canada*.
- Liao, L. A. (2006). A flow theory perspective on learner motivation and behavior in distance education. *Distance Education*, 27(1), 45–62. <http://dx.doi.org/10.1080/01587910600653215>.
- Ma, Y., & Williams, D. (2011). Designing an Electronic Educational Game to Facilitate Immersion and Flow. In *Paper presented at the AERA Annual Meeting, New Orleans, LA*.
- MacCallum, R. C., Wildaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4(1), 84–89. <http://dx.doi.org/10.1037/1082-989X.4.1.84>.
- Malone, T. W., & Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. In R. E. Snow & M. J. Farr (Eds.), *Aptitude learning and instruction* (Vol. 3, pp. 223–253). Mahwah, NJ: Erlbaum.
- Mandryk, R. L., Inkpen, K. M., & Calvert, T. W. (2006). Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour and Information Technology*, 25(2), 141–158. <http://dx.doi.org/10.1080/01449290500331156>.
- Marsh, H. W., & Jackson, S. A. (1999). Flow experience in sport: Construct validation of multidimensional, hierarchical state and trait responses. *Structural Equation Modeling – A Multidisciplinary Journal*, 6(4), 343–371. <http://dx.doi.org/10.1080/10705519909540140>.
- Mayer, R. E. (2003). Elements of a science of e-learning. *Journal of Educational Computing Research*, 29(3), 297–313.
- Mayo, M. J. (2009). Video games: A route to large-scale STEM education? *Science*, 323(5910), 79–82.
- O'Brien, H. L. (2010). The influence of hedonic and utilitarian motivations on user engagement: The case of online shopping experiences. *Interacting with Computers*, 22(5), 344–352. <http://dx.doi.org/10.1016/j.intcom.2010.04.001>.
- O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, 59(6), 938–955. <http://dx.doi.org/10.1002/asi.20801>.
- O'Brien, H. L., & Toms, E. G. (2010). The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology*, 61(1), 50–69. <http://dx.doi.org/10.1002/asi.21229>.
- O'Brien, H. L., & Toms, E. G. (2012). Examining the generalizability of the User Engagement Scale (UES) in exploratory search. *Information Processing and Management*, 49(5), 1092–1107. <http://dx.doi.org/10.1016/j.ipm.2012.08.005>.
- Picard, R. W. (2010). Affective computing: From laughter to IEEE. *IEEE Transactions on Affective Computing*, 1(1), 11–17. <http://dx.doi.org/10.1109/t-affc.2010.10>.
- Przybylski, A. K., Rigby, C. S., & Ryan, R. M. (2010). A motivational model of video game engagement. *Review of General Psychology*, 14(2), 154–166. <http://dx.doi.org/10.1037/a0019440>.
- Rieber, L. P. (1996). Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educational Technology Research and Development*, 44(2), 43–58.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78.
- Sharek, D. (2009). Block Walk (Version 1.0). Raleigh, NC.
- Sharek, D. (2012). Investigating real-time predictors of engagement: Implications for adaptive video games and online training. (Ph.D.), North Carolina State University, Raleigh, NC.
- Sherry, J. L. (2004). Flow and media enjoyment. *Communication Theory*, 14(4), 328–347. <http://dx.doi.org/10.1111/j.1468-2885.2004.tb00318.x>.
- Shneiderman, B. (1998). *Designing the user interface: Strategies for effective human-computer interaction* (3rd ed.). Reading, MA: Addison-Wesley.
- Skelly, T., Fries, K., Linnett, B., Nass, C., & Reeves, B. (1994). Seductive interfaces: Satisfying a mass audience. In *Paper presented at the Conference Companion on Human Factors in Computing Systems, Boston, Massachusetts, USA*.
- Stephenson, W. (1967). *The play theory of mass communication*. Chicago: University of Chicago Press, pp. 45–65.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson.
- Turkay, S. (2012). User experiences with avatar customization in second life and lord of the rings online. In *Paper presented at the fourth annual teachers college educational technology conference, New York, NY*.
- Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences*, 39(2), 273–315.
- Vorderer, P., Klimmt, C., & Ritterfeld, U. (2004). Enjoyment: At the heart of media entertainment. *Communication Theory*, 14(4), 388–408. <http://dx.doi.org/10.1111/j.1468-2885.2004.tb00321.x>.
- Weibel, D., Wissmath, B., Habegger, S., Steiner, Y., & Groner, R. (2008). Playing online games against computer- vs. human-controlled opponents: Effects on presence, flow, and enjoyment. *Computers in Human Behavior*, 24(5), 2274–2291. <http://dx.doi.org/10.1016/j.chb.2007.11.002>.
- Whitton, N. (2011). Game engagement theory and adult learning. *Simulation and Gaming*, 42(5), 596–609. <http://dx.doi.org/10.1177/1046878110378587>.