Exploring Behavioral Expressions of Player Experience in Digital Games

Wouter van den Hoogen
Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven,
The Netherlands
W.M.v.d.Hoogen@tue.nl

Wijnand IJsselsteijn
Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven,
The Netherlands
W.A.IJsselsteijn.nl@tue.nl

Yvonne de Kort
Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven,
The Netherlands
Y.A.W.d.Kort@tue.nl

Abstract

This paper describes a first exploration of human motor behavior that may be associated with player experiences in digital games. Evidence from literature suggests that patterns in pressure and postural movement data may be indicative for experiences such as interest, arousal, frustration and boredom. In the current study we explore the relation between behavioral measures and people's emotional experience during game play. Results from the study presented in this paper indicate that the intensity of people's actions (e.g. pressure exerted on the mouse) and bodily movement relates to several experiences during game-play, including frustration. However, the results show that these behavioral measures do not exclusively relate to one specific experience. Rather, the results imply these behavioral measures to relate to the level of arousal and level of dominance felt during game-play. From these results it is evident that behavioral measures have a clear application potential. This study presents a starting point in the development of a set of behavior-based measures of player experiences. Establishing sensitivity and validity of such measures can be regarded as the necessary first step in the process of creating an emotionally adaptive game.

1. Introduction

One of the main challenges facing the digital games research community is the development of a coherent and fine-grained set of methods and tools that enable the measurement of entertainment experiences in a sensitive, reliable and valid manner. Measures that capture users' emotions and experiences during gameplay will substantially enhance our understanding of game elements that are particularly engaging and motivating. This will likely aid theory development by allowing a much more direct coupling between specific game design patterns [1] and player experiences. Moreover, understanding gameplay at its base level will allow game designers to introduce those design elements in a game which are known to elicit the most engaging experiences, based on an understanding of what the player will be experiencing at each point in the game. Eventually, the output of continuous measures of player experiences may become real-time input to the game engine, allowing the game's artificial intelligence to adjust to the player's affective or cognitive state at any point during gameplay.

It should be noted that in the large body of literature on media reception and reaction processes, the behavioral impact of media is usually discussed in terms of how media affect behavioral tendencies after episodes of media exposure. For example, a significant body of digital games research is looking at potential associations between exposure to violent games and the development and manifestation of antisocial (e.g., aggressive) behaviors [6]. However, when we refer to behavioral responses in the current paper, we are referring to naturally occurring physical and social behaviors as they are exhibited during an episode of gameplay, as a direct response to unfolding game events and/or social interactions among multiple game participants.

The current paper sets out to describe a first exploration of behavioral expressions that could serve as real-time indicators of experiences related to playing digital games. In this paper, we focus primarily on pressure patterns exerted on a physical control device, and postural responses. Based on this exploration, we present our progress in developing a set of behavior-based measures of such player experiences and their application in an

978-1-4244-2154-1/08/$25.00 ©2008 IE

experimental study.

1.1. Flow, frustration and boredom

Csikszentmihalyi [4,5] studied what makes experiences enjoyable to people. He was interested in people's inner states while pursuing activities that are difficult, yet appear to be intrinsically motivating, that is, contain rewards in themselves – chess, rock climbing, dance, sports. In later studies, he investigated ordinary people in their everyday lives, asking them to describe their experiences when they were living life at its fullest, and were engaged in pleasurable activities. He discovered that central to all these experiences was a psychological state he called flow, an optimal state of enjoyment where people are completely absorbed in the activity. Flow is a state where someone's skills are well balanced with the challenges posed by a task. It is characterized by a deep concentration on the task at hand, a perceived sense of control over actions, a loss of preoccupation with self, and transformation of one's sense of time.

Flow certainly sounds familiar to frequent players of computer games. Digital games provide players with an activity that is goal-directed, challenging and requiring skill. Most games offer immediate feedback on distance and progress towards the (sub)goals, through, for instance, score keeping, status information (e.g., a health indicator), or direct in-game feedback. When a game is effective, the player's mind can enter an almost trance-like state in which the player is completely focused on playing the game, and everything else seems to fade away - a loss of awareness of one's self, one's surroundings, and time. It is the experience that is strongly connected to what gamers and game reviewers commonly refer to as the 'gameplay' of a game, i.e., the somewhat ambiguous term describing a holistic gaming experience, based on a fluent interaction with all active gaming elements, the progression of challenges offered, and the ability of a game to continuously command the attention of a player.

Sweetser and Wyeth [19] have adopted and extended Csikszentmihalyi's conceptualization of flow in their 'GameFlow' model of player enjoyment, formulating a set of useful design criteria for achieving enjoyment in electronic games – see also [8]. Csikszentmihalyi's original work on flow suggests that these peak experiences are quite rare – the exception rather than the rule. Nevertheless, the flow model of game enjoyment clearly illustrates the importance of providing an appropriate match between the challenges posed and the player's skill level. The flow experience can easily break down when the player's skills systematically outpace the challenges the game can offer (leading to boredom) or when game challenges become overwhelming in light of the available skills (resulting in frustration). Challenge is probably one of the most important aspects of good game design, and adjusting the challenge level to accommodate the broadest possible audience in terms of player motivation, experience and skill is a major challenge for current game designers.

Being able to detect frustration and boredom is of importance as indicators of when a person is not experiencing flow, but also, and perhaps more interestingly, because successful games strike a balance between positive and negative emotions (see, e.g., [16]). This is in line with the view that games are often being designed with the aim to develop a negative emotion in the face of challenge, only to be followed by a positive emotional peak when the challenge is overcome [9]. In sum, behavioral indicators of involvement or interest are required, as well as indicators of both boredom and frustration.

1.2. Behavioral expression of player experiences

Behavioral expressions of subjective states are well known to both lay-people and scientists alike. A host of observable and expressive physical behaviors are associated with emotional states. We tend to smile at something funny, move towards something or somebody we like, jump up when startled, hide our heads when scared, or make strong gestures when frustrated. There are a number of behavioral responses where the human motor system may potentially act as a carrier for the player experiences discussed previously.

Mota & Picard [14] demonstrated that postural patterns can be indicative of learner interest. They developed a system to recognize postural patterns and associated affective states in real time, in an unobtrusive way, from a set of pressure sensors on a chair. Their system is reportedly able to detect, with an average accuracy of 87.6%, when a child is interested, or is starting to take frequent breaks and looking bored. Thus, the dynamics of postures can distinguish with significant reliability between affective states of high interest, low interest and boredom, all of which are of relevance to a gaming situation as well.

Clynes [2,3] investigated the patterns of motor output of people asked to deliberately express certain emotions through the motor channel (usually a finger pressing on a measuring surface he dubbed the 'sentograph'). He found that there are distinguishable, stable patterns of pressure and deflection for emotions such as anger, hate, grief, love, and joy, transcending barriers of culture and language [2]. Support for Clynes' original findings has been varied. Trussoni, O'Malley and Barton [21] failed to replicate Clynes' findings using an improved version of the sentograph. Although they did find distinguishable patterns associated with certain emotions, a significant correlation with Clynes' original sentograms [2] was absent, throwing doubt on the universality of sentic patterns. Hama and Tsuda [7], on the other hand, did find support for the characteristic waveform patterns associated with 'sadness' (long duration of pressure) and 'anger' (strong intensity of pressure). Moreover, in their first experiment, Hama and Tsuda did not inform participants that they were interested in measuring emotions, which raises the interesting possibility that identifiable pressure patterns may be associated with spontaneously generated motor expression of emotions. In particular, the sentic expression of anger is of interest as a potential indicator of gamer frustration.

Research by Mentis and Gay [13] and Park, Zhu, McLaughlin & Jin [15] provide evidence that the force people apply to interface devices can be interpreted as an indicator of negative arousal. Mentis and Gay [13] asked a small number of participants to complete several tasks on a word processor. Later, participants were asked to indicate whether and when they experienced a frustrating event. Their results suggest that higher pressure on the touchpad is associated with a frustrating event. Building on these findings, Park et al. [15] manipulated frustration by asking participants to complete an impossible LEGO assembly task. The instructions for the task and optional online help were presented on a laptop computer, where the pressure exerted on the touchpad was measured. Results indicated that more pressure was exerted on the interface device when participants were encountering problems. Additionally, pressure patterns also correlated with facial expressions showing negative affect, thereby providing evidence that the pressure exerted was indeed related to frustration rather than mere arousal.

Focusing on digital games, Sykes and Brown [20] have investigated the mean pressure exerted by players on a gamepad's button as the difficulty level of a game (Space Invaders) was increased from easy to medium to hard. Their results show that buttons on the gamepad were pressed significantly harder in the hard condition than in either the easy or the medium condition. Although the increase in pressure on the gamepad can be assumed to be associated with higher arousal, Sykes and Brown did not determine whether this arousal was positively or negatively valenced, while both states could plausibly occur in a digital game setting. Notwithstanding this limitation, Sykes and Brown [20] successfully demonstrated that a fairly straightforward behavioral measure such as hand or finger pressure exerted on a button can already be informative about the level of user arousal in gaming situations. In addition, given its relative simplicity, this measure has the potential to be analyzed in real-time and be used to adaptively influence the game dynamics.

2. Linking behavior to player experience

From the literature it is evident that behavioral patterns are likely to be informative for the real-time measurement of player experiences. From a methodological point of view, there are several advantages associated with employing behavioral measures as an indicator of player experiences. First, they are relatively free from subjective bias, because they are generally not under users' conscious control, nor do they require specific instructions from an experimenter (e.g., "please hit the button harder as you get more frustrated") – they occur spontaneously. Secondly, when measured in an unobtrusive fashion, they do not disrupt the player experience. Third, they are time-continuous measures, that is, they are collected as the experience is unfolding, and are as such not reliant on memory or introspection on the part of the participant (unlike self-report measures). Finally, a number of these measures, such as a pressure-sensitive gamepad, could realistically be integrated with existing game technologies. This is a clear advantage when these measures are to be integrated in commercial games, where specialist peripheral hardware will only scarcely be adopted.

In the current study, we want to explore a number of behavioral measures in relation to player experiences. The aim of such an exploration is twofold. First, we need to establish which behavioral measures are sensitive to variations in game dynamics. Second, we need to find out in what way behavioral measures are correlated to player experiences, thereby establishing a potential connection between objective measurements and subjective experience. Behavioral indicators that are demonstrated to be both sensitive to experimental manipulations and sensibly related to player experiences can subsequently be deployed in closing the loop between the player and the game. That is, successful behavioral indicators of player experience can be used as real-time input data to the game engine, dynamically adapting the game to the player's experiential state. The current study should thus be regarded as the necessary first step in the process of creating an emotionally adaptive game, establishing sensitivity and validity of behavioral indicators of player experiences.

In an attempt to link measurable behavior such as postural movements and pressure patterns to people's emotional states during digital game play, we have recently developed several real-time behavior measurement systems, including a pressure-sensitive chair, inspired on the work of Mota and Picard [14], and a pressure sensitive mouse and keyboard. Although we have reviewed and tried various off-the-shelf solutions, including VR pressure-sensitive gloves, our fairly straightforward, customized measures allow for more sensitive measurement of various bodily responses, are not overly obtrusive, and can be easily integrated with existing gaming devices. Moreover, the combination of *multiple* behavioral indicators can reduce uncertainty or ambiguity associated with a single indicator, resulting in increased robustness and wider applicability of the total set of measures. Limitations particular to one measure may be overcome or compensated by using corroborating evidence emerging from another measure.

In the study reported in this paper we therefore decided to use multiple behavior measurement systems in conjunction with self report measures of people's game-play. Within this study we have used customized levels of a digital game (Half Life 2) to induce boredom, enjoyment, and frustration. By inducing these player experiences, we do not need to infer these states, nor wait for their spontaneous occurrence. Moreover, such a manipulation is expected to result in much needed variation in types of experiences. This will allow us to more reliably associate behavioral response patterns with affective states (see also [17]).

3. Method

The experiment was conducted in the Game Experience Lab at Eindhoven University of Technology. The first person shooter game Half Life 2 was modified such that game difficulty was either easy, moderate, or hard, according to a within groups design. After each level participants filled in a questionnaire including the self report measures aimed to measure player experience. The game was played on a Dell XPS PC equipped to cope with the demands of the game and was connected to a 20" TFT-screen.

3.1. Participants

Thirty-two participants (five females) aged between 17 and 46 (Mage = 22.42 years, SD = 5.57 years) took part in the experiment. All participants at least occasionally played first person shooter (FPS) games, but a substantial part consisted of more frequent players. Participants received 10€ for their time.

3.2. Procedure

Upon entering the lab, participants were welcomed by the experiment leaders. The experiment leaders gave a brief overview of the progression of the experiment. Participants signed the consent form (allowing video observations and psychophysiological measures to be taken), were seated at a desk where the game-PC was installed, and were connected to psychophysiological sensors and an accelerometer. After reading brief instructions related to the use of the controls in the game, participants played the three customized levels of the FPS game Half Life 2. After each level, participants rated their experiences during game-play on a range of self-report measures administered on a separate laptop PC. The order in which the levels were played was counterbalanced. Participants were given ten minutes to play each of the levels with the exception of the easy level. Because more experienced players usually finished the easy level in less than ten minutes, we fixed the playtime for this level at eight minutes. At the end of the session, participants were paid, debriefed, and thanked for their participation.

3.3. Measures

The study was designed to relate behavioral responses to self reported experiences during game play. Consequently, during the study we measured both people's self reported experience of each level played and measured their behavior using a range of behavioral measurement tools. The measures are described in more detail below.

3.3.1 Self report measures

Self report measures used in the study included the Self Assessment Manikin (SAM) scale [10;18] and the in-game version of the Game Experience Questionnaire [22] recently developed by the Game Experience Lab at Eindhoven University of Technology. Further, we included a manipulation check for the level of difficulty.

3.3.1.1 SAM-scale

The SAM scale is a visual self report scale developed by Lang [10] and based on Mehrabian and Russell's [12] Pleasure-Arousal-Dominance (PAD) theory. The SAM-scale visualizes the three PAD-dimensions. Each dimension is depicted through a set of five graphic figures (manikins) and for every dimension respondents have to indicate which figure corresponds best with their feelings on a nine point scale. The first dimension P (displeasure/pleasure) ranges from extreme sadness to extreme happiness. The second dimension A (non-arousal/arousal) ranges from very calm or bored to extremely stimulated. The third dimension D (submissiveness/dominance) ranges from a feeling of being controlled or dominated to a feeling of total control. Additionally, we included a SAM-based measure of presence developed by Schneider et al. [18] as a fourth emotion dimension that possibly applies to digital game experience. This dimension ranges from a feeling of total presence to a feeling of total absence. For each SAM dimension we asked participants to indicate, on a 9-point scale listed below the graphical presentation, which manikin corresponded with their experiences during game-play. Scale values ranged from -4 to 4, with ascending scores corresponding to higher pleasure, higher arousal, higher dominance and lower presence ratings.

3.3.1.2 In-Game GEQ (iGEQ)

After each level we administered the in-game Game Experience Questionnaire (iGEQ) consisting of seven dimensions with two items per dimension. These dimensions were: Positive affect *(I felt content, I felt good*), Boredom (*I found it tiresome, I felt bored*), Frustration (*I felt irritable, I felt frustrated*), Flow (*I felt completely absorbed, I forgot everything around me*), Challenge (*I felt stimulated, I felt challenged*), Immersion (*I was interested in the game's story, I found it impressive*), and Competence (*I felt skilful, I felt successful*). All GEQ items are measured by means of five point intensity scales with points anchored at not at all (0), slightly (1), moderately (2), fairly (3), extremely (4). For our analyses, we used the mean value of the two items per dimension. We used the iGEQ, the shorter in-game version of the GEQ, because we did not want to interrupt participants too long between the different levels of game-play.

3.3.1.3 Manipulation check

The manipulation check included one five point bi-polar statement stating "How easy or difficult did you find it to play the level?" ranging from -2 (too easy to play) via 0 (optimal to play) to 2 (too difficult to play).

3.3.2 Behavioral measures

During the game-play we measured people's movement on the chair they were sitting on, measured the movement of their upper body by means of an accelerometer, and measured the force they applied to the mouse. Each of these measures is shortly explained below.

3.3.2.1 Accelerometer

For each participant, an accelerometer was attached to the back, at the base of the neck, to automatically capture movement of the upper body. The accelerometer used was a Phidgets 3 axis version measuring tilt on the x, y, and z-axes, and acceleration to a maximum of 3Gs, which is more than enough for the expected movement of the participants during game play. For the analyses we used the accelerometer data converged over all axes (square root of the sum of the squared values for each of the three other axes). Subtracting the mean value across all levels from the individual data values and calculating the absolute value resulted in a metric representing the acceleration as a function of movement in any direction. In addition to the maximum value per level, these values were averaged per level providing an indication of the average movement during each level.

3.3.2.2 Pressure sensitive chair

A second automatic indicator of movement was recorded via a pressure sensitive chair. Sitting position and the number of shifts in position are potential indicators of boredom and of interest. In addition to observed and coded sitting position (forward-backward movement) using the video streams, we also employed a custom-built posture-sensitive chair using force-sensitive sensors built into the legs of the chair. This allowed real-time measurement of the forward-backward and sideways movements of the participant during game-play. The sensors used were TekScan pressure sensitive sensors designed to measure up to 25Lbs (approx. 11.3 Kg) of force applied to them (for an image of the chair and the measuring system see Figure 1).

Figure 1: Pressure sensitive chair used for the measurement of people's changes in sitting position.

For the purposes of the current study we calculated the maximum range of forward-backward movement on the chair by subtracting the minimum value (most backward position) from the maximum value (most forward position). This measure was calculated for each of the levels played. As there are likely large individual differences in the rate of movement we applied a range correction to the measures. That is, the values of the range of movement for each level were divided by the maximum range across all levels for that individual. This procedure was used as this is advised for the use of galvanic skin response (GSR) data [11] which has similar properties and dependencies on individual differences to our automatically captured behavioral measures. Additionally, it neutralizes potential differences in sensitivity of the pressure sensors (e.g. due to differences in weight of the participants), and allows comparison across individuals.

3.3.2.3 Pressure sensitive mouse

The mouse was equipped with two Flexiforce sensor designed to measure up to 1Lbs [approx. 453.6 grams] of force applied to them, mounted on top of two buttons. To increase the likelihood that the participants would press on the sensors when operating the mouse the paddles were reduced in size and the sensors were topped with a small rubber patch. This patch raised the surface of the sensor over the rest of the paddle and discriminated the surface texture of the paddles. The patch thus naturally invited people to keep their fingers on top of the sensors (see Figure 2 for a view of the augmented mouse).

Figure 2: Pressure sensitive mouse measuring force applied to the mouse during game-play.

The mouse pressure data were recorded continuously allowing for synchronization of the force on the input devices with discrete in-game events. The data can be aggregated over lengths of time, e.g., complete sessions. This provides opportunities for event-based analyses, and correlation analyses with self-report measures. For analyses of the force applied to the mouse, two measures were constructed. The first measure was constructed using the maximum value of force applied to the left mouse button per level. As with the chair we applied a range correction to the values. That is, the maximum value of each level was divided by the overall maximum value of force during game-play. Again this was done to reduce individual differences in the force people apply on an interface device, allowing comparison between participants. The second measure constructed was the average force applied to the mouse based on the maximum force per event, thus excluding all values between the onset and end of the mouse press other than the maximum force. Again like the maximum mouse force, the mean force was based on the range corrected values.

4. Results

First the results of the manipulation check will be presented, followed by the results of the behavioral measures. Although the results of the self-report measures will be reported elsewhere, the correlation between the self-report measures and the behavioral measures are reported in this paper, since this provides an indication of the validity of the behavioral measures.

4.1. Manipulation Check

The three levels used in this study were designed to represent an easy, a challenging, and a hard level in terms of difficulty, ideally inducing boredom, flow/ enjoyment, and frustration. In order to establish the effect of the manipulation we conducted a repeated measures ANOVA on the one-item manipulation check '*How easy or difficult did you find it to play the level?*'. This analysis showed significant differences between each of the three difficulty levels in the expected directions ($F(2,30)=120.77$, $p<.001$). The easy level was rated as the "easiest" (M= -1.47, SD= 0.84) followed by the moderate level (M= -0.5, SD= 0.95), and the hard level (M=1.09, SD= 0.73) as the most difficult level to play. This result thus provides initial confirmation that the difficulty manipulation was effective.

4.2. Behavioral measures

4.2.1 Accelerometer

Using a repeated measures ANOVA we analyzed the effects of the level of difficulty on both the maximum and mean scores acquired. A Greenhouse Geisser correction for the repeated measures ANOVA was used, correcting for violations of sphericity. The results indicate that there is no difference in the maximum accelerometer data between the levels (see Table 1). The mean accelerometer value did, however, differ between the levels ($F(1.61,29.39)=10.69$, $p<.001$). The mean value for the hard level proved to be highest and significantly different from both the moderate and the easy level. This implies that in the hard level participants, on average, moved more strongly than in the other levels.

4.2.2 Sitting position

The second behavioral indicator of player movement was acquired via the sensors in the pressure sensitive chair. This indicator takes into account not only the movement of the upper part of the body, but rather the center of gravity of the body as a whole. Sitting position was analyzed using the corrected range from the forward-backward position on the chair. The results from the repeated measures ANOVA showed the range in movement on the chair to differ significantly between the levels ($F(2,29)=5.52$, $p=.006$) with the hard level having the highest score and differing from the easy and moderately difficult levels (see Table 1). In line with the accelerometer results, this implies that movement was strongest in the most difficult level.

4.2.3 Mouse pressure

Sensors on the left mouse buttons measured the force with which players made each mouse click. The force applied to the mouse was analyzed using both the maximum mouse force and the mean mouse peak force. Both indicators increased with the difficulty of the game level (see Table 1). The maximum mouse force differed significantly between the levels ($F(2,29)=11.72$, $p<.001$), with the hard level differing significantly from both the easy and moderately difficult levels. As for the maximum mouse force, mean mouse force was highest in the hard level. The effect was however only marginally significant, with the easy and hard level differing from each other. The results show that, on average, people applied most force on the mouse in the difficult level. Similar to the accelerometer and chair results, this result implies that the behavior was again most intense in the most difficult level.

TABLE 1: MEANS OF BEHAVIORAL MEASURES PER LEVEL OF DIFFICULTY WITH STANDARD DEVIATIONS IN PARENTHESES
(† = MARGINALLY SIGNIFICANT, *P<.05, ** P<.01, *** P<.001)
((a,b,c): DIFFERS SIGNIFICANTLY (PAIRWISE COMPARISONS, P < .05) FROM EASY LEVEL(a), MODERATE LEVEL(b), HARD LEVEL(c))

	Easy	Moderate	Hard
Mean Accelerometer ***	0.0036 ^{c} (.0008)	0.0037 ^{c} (.0008)	0.0041 ^{a,b} (0.0012)
Maximum Accelerometer	0.13 (0.18)	0.13 (0.09)	0.13 (0.07)
Chair Movement **	0.59 ^{c} (0.31)	0.56 ^{c} (0.32)	0.81 ^{a,b} (0.28)
Mean Mouse Force †	0.11 ^{c} (0.059)	0.12 (0.067)	0.13 ^{a} (0.069)
Maximum Mouse Force ***	0.45 ^{c} (0.31)	0.54 ^{c} (0.32)	0.84 ^{a,b} (0.28)

4.3. Correlations between self report measures and behavioral measures

The results demonstrate that the behavioral measures used in this study (mouse force, movement on a chair, and upper body movement) related to the level of difficulty. More importantly, with the exception of the maximum accelerometer value they were highest when the level of difficulty was highest, in line with previously reported findings (e.g. [20]). However, as earlier research makes clear, exactly what these measures (e.g. mouse force) indicate is unclear. Although mouse force has previously been associated with frustration [13], in gaming this measure may signify both pleasurable challenge as well as frustration. In order to connect the behavioral measures to player experience, we included the iGEQ, and the SAM. In this section we present the correlations between the behavioral measures found to be sensitive to the manipulation of difficulty and the self report measures.

For thoroughly exploring correlations between variables, sufficient variation is needed. Since the experimental levels were explicitly created to induce a specific experience, variance within each level was only modest. For this reason we restructured the data such that the different experimental levels were treated as separate cases, creating three rows of data for each participant. By exploring correlations across levels we created variation in the different measures enabling us to report reliable conclusions about how the self report measures are related to the behavioral measures.

TABLE 2: CORRELATIONS BETWEEN BEHAVIORAL MEASURES AND iGEQ.
(† = MARGINALLY SIGNIFICANT, * P<.05, ** P<.01, ***P<.001)

	Maximum Mouse Force	Mean Accelerometer	Chair Movement
Immersion	.063	-.110	.051
Competence	-.392 ***	-.197 †	-.289 **
Neg. Affect (boredom)	-.262 *	-.052	-.226 *
Flow	.205 *	-.048	.199 †
Frustration	.335 ***	.148	.356 ***
Challenge	.399 ***	.302 **	.252 *
Pos. Affect	-.141	.057	-.240 *

From Table 2 it is evident that the behavioral measures are correlated with multiple dimensions of the iGEQ, rather than only one as previous research suggests [13;15]. The results thus imply the behavioral measure to be related to more than only one specific emotion (such as frustration). Most notably, there is a large overlap in direction and magnitude of the correlations with the iGEQ dimensions of both the Maximum Mouse Force and Chair Movement.

Additionally, the behavioral measures are negatively correlated with items that can be interpreted as being low arousal experiences (Positive Affect, Boredom, and Competence), while they are positively correlated with items signaling higher arousal states (Frustration, Challenge, and to a lesser extent Flow). This suggests that the intensity of behavior (chair movement and pressure on the interface device) is an indicator of arousal as underlying physiological state of the person playing the game. Correlations of the behavioral measures with the SAM are indeed consistent with this interpretation. As can be seen in Table 3 the behavioral measures are positively correlated with the arousal dimension and negatively correlated with the dominance dimension. Indeed in the context of game-play it seems that these two are to some extent each other counterparts. As people feel they are more dominated by the game (i.e. lose control) they will likely get more aroused through this challenge.

In sum, the correlations of the behavioral measures with the SAM and iGEQ dimensions show that the behavioral measures are likely indicators of arousal, more so then they can be interpreted as indicators of one specific emotion or experience.

TABLE 3: CORRELATIONS BETWEEN BEHAVIORAL MEASURES AND SAM.DIMENSIONS
(† = MARGINALLY SIGNIFICANT, * P<.05, ** P<.01, ***P<.001)

	Maximum Mouse Force	Mean Accelerometer	Chair Movement
Pleasure	-.150	.073	-.160
Arousal	.189 †	.219 *	.222 *
Dominance	-.301 **	-.236 *	-.273 **
Presence	.053	.061	-.038

5. Conclusions

In the study presented in this paper we investigated the potential for multiple behavioral measures as indicators of people's game-play experience. Using an experiment in which difficulty level was manipulated, we have found automatically captured body movement and pressure on the interface device to be highest in the most difficult level. These findings are in line with our hypotheses, and support and extend earlier findings by Sykes and Brown [20] who found that more pressure was exerted on a gamepad's button as the difficulty level increased. However, Sykes and Brown did not take any self-report measures to help interpret pressure as a measure of player experience. Our results suggest there is no easy one-to-one relation with frustration or enjoyment. Rather, we find that an increase in arousal, be it through increased frustration or challenge, results in a higher level of pressure exerted. A similar pattern emerges for the measures of postural movement.

Combined, our findings suggest that measures of movement and pressure mainly serve as indicators of people's level of arousal. The intensity of action has been found to relate to arousal states: i.e. they were highest when the level was most difficult, correlated positively with the high arousal experiences, were negatively correlated with low arousal experiences, and were positively correlated with SAM arousal and dominance scales. Our findings do not support suggestions made in previous research that pressure exerted on a mouse (or touchpad) is *exclusively* associated with the experience of frustration [13, 15]. It is important to note that such results have been obtained in productivity oriented tasks (word processing task, LEGO assembly task), where frustrating events are explicitly included (e.g., a task that is impossible to complete), and positive challenge is lacking. In contrast, a digital game such as Half Life 2 allows for a more varied spectrum of challenges, some of which add to the excitement of the game, others leading to frustration. Thus, more force applied to the interface device cannot be simply translated to higher levels of frustration. The study presented in this paper is a first exploring the relation between multiple behavioral measures and multiple self reported dimensions of game-play. Importantly, our research show not *only* mouse force to correlate with multiple experiences. These relations are evident for chair movement, and to a lesser extent upper body movement as well. Further, the relations appear to be consistent across the behavioral measures included in our experiment, in all cases the intensity of the behavior appears to relate to the level of arousal that players experience. The consistency of these findings bode well for behavioral indicators as potential input data to game engines.

Having made a first step in determining the sensitivity and validity of a number of behavioral measures, we will next turn our attention to determining whether these relations are stable over time (test-retest reliability), as well generalisable across different games and gaming genres. Moreover, although we have used time-continuous measures, we have analysed them in aggregate form (i.e., means across levels and players). In order to firmly establish whether such measures can be useful as input to emotionally adaptive games, we need to establish sensitivity of the measures at an individual level, and across much shorter time-spans (in the order of seconds rather than minutes). Further, analyses of our current rich dataset, as well as new experiments, are expected to throw light on this issue in the near future.

Acknowledgments

Thanks to Martin Boschman for technical assistance and co-design of the measurement systems. Funding from the EC through the IST Framework 6 FUGA (http://project.hkkk.fi/fuga/) and Games@Large projects (http://www.gamesatlarge.eu/) is gratefully acknowledged.

References

[1] Björk, S. and Holopainen, J. (2004). Patterns In Game Design. Charles River Media.

[2] Clynes, J.M. (1973). Sentics: Biocybernetics of emotional communication. Annals of the New York Academy of Science, 220, 55-131.

[3] Clynes, J.M. (1977). Sentics: The Touch of the Emotions. Anchor Press/Doubleday, 1977.

[4] Csikszentmihalyi, M. (1975). Beyond Boredom and Anxiety. San Fransisco: Jossey-Bass.

[5] Csikszentmihalyi, M. (1990). Flow. The Psychology of Optimal Experience. New York: Harper & Row.

[6] Gunter, B. (2005). Psychological effects of video games. In: J. Raessens & J. Goldstein (eds.), Handbook of Computer Game Studies. Cambridge, MA: MIT Press, pp. 145-160.

[7] Hama, H. and Tsuda, K. (1990). Finger-pressure waveforms measured on Clynes' Sentograph distinguish among emotions. Perceptual and Motor Skills, 70, 371-376.

[8] Holt, R. and Mitterer, J. (2000). Examining video game immersion as a flow state. Paper presented at the 108th Annual Psychological Association, Washington, DC.

[9] Keeker, K., Pagulayan, R., Sykes, J., and Lazzaro, N. (2004). The untapped world of video games. ACM CHI 2004, 1610-1611.

[10] Lang, P.J. (1980). Behavioral Treatment and Bio-Behavioral Assessment: Computer Applications. In Technology in mental health care delivery systems, Sidowski, Joseph B., James H. Johnson, and Thomas A. Williams (Eds.). (119-37). Norwood, NJ: Ablex Publishing.

[11] Lykken, D.T., & Venables, P.H. (1971) Direct measurement of skin conductance: A proposal for standardization. Psychophysiology, 8 (5), 656-672.

[12] Mehrabian, A., & Russell, J.A. (1974). An Approach to Environmental Psychology. Cambridge, MA: The MIT Press.

[13] Mentis, H.M. and Gay, G.K. (2002). Using touchpad pressure to detect negative affect. Proceedings of Fourth IEEE International Conference on Multimodal Interfaces 2002, 406 - 410

[14] Mota, S. and Picard, R.W. (2003). Automated Posture Analysis for Detecting Learner's Interest Level. Workshop on Computer Vision and Pattern Recognition for Human-Computer Interaction (CVPR HCI). Available: http://affect.media.mit.edu/pdfs/03.mota-picard.pdf

[15] Park, N., Zhu, W., Jung, Y., McLaughlin, M., and Jin, S., (2005). Utility of haptic data in recognition of user state. Proceedings of HCI International 11. Lawrence Erlbaum Associates. Available: http://imsc.usc.edu/haptics/paper/manuscript_hcii2005_final.pdf

[16] Ravaja, N., Salminen, M., Holopainen, J., Saari, T., Laarni, J. and Järvinen, A. (2004) Emotional response patterns and sense of presence during video games: potential criterion variables for game design. Proceedings of the Third Nordic Conference on Human-Computer Interaction, 339-347.

[17] Scheirer, J., Fernandez, R., Klein, J., and Picard, R.W. (2002). Frustrating the user on purpose: A step toward building an affective computer. Interacting with Computers, 14, 2, 93-118.

[18] Schneider, E.F., Lang, A., Shin, M., & Bradley, S.D. (2004). Death with a story: How story impacts emotional, motivational, and physiological responses to first-person shooter video games. Human Communication Research, 30 (1), 361-375.

[19] Sweetser, P and Wyeth, P. (2005). GameFlow: A model for evaluating player enjoyment in games. ACM Computers in Entertainment, 3 (3), 1-24.

[20] Sykes, J. and Brown, S. (2003). Affective gaming. Measuring emotion through the gamepad. ACM CHI 2003, 732-733.

[21] Trussoni, S.J., O'Malley, A., and Barton, A. (1988). Human emotion communication by touch: A modified replication of an experiment by Manfred Clynes. Perceptual and Motor Skills, 66, 419-424.

[22] IJsselsteijn, W.A., de Kort, Y.A.W. & Poels, K. (in preparation). The Game Experience Questionnaire: Development of a self-report measure to assess the psychological impact of digital games. Manuscript in preparation.