Experiencing BCI Control in a Popular Computer Game

Bram van de Laar, Hayrettin Gürkök, Danny Plass-Oude Bos, Mannes Poel, and Anton Nijholt

***Abstract*—Brain–computer interfaces (BCIs) are not only being developed to aid disabled individuals with motor substitution, motor recovery, and novel communication possibilities, but also as a modality for healthy users in entertainment and gaming. This study investigates whether the incorporation of a BCI in the popular game *World of Warcraft (WoW)* has effects on the user experience. A BCI control channel based on parietal alpha band power is used to control the shape and function of the avatar in the game. In the experiment, participants ($n = 42$), a mix of experienced and inexperienced *WoW* players, played with and without the use of BCI in a within-subjects design. Participants themselves could indicate when they wanted to stop playing. Actual and estimated duration was recorded and questionnaires on presence and control were administered. Afterwards, oral interviews were taken. No difference in actual duration was found between conditions. Results indicate that the difference between estimated and actual duration was not related to user experience but was person specific. When using a BCI, control and involvement were rated lower. But BCI control did not significantly decrease fun. During interviews, experienced players stated that they saw potential in the application of BCIs in games with complex interfaces such as *WoW*. This study suggests that BCI as an additional control can be as much fun and natural to use as keyboard/mouse control, even if the amount of control is limited.**

***Index Terms*—Brain–computer interface (BCI), games, human factors, presence, user experience.**

I. Introduction

BRAIN–COMPUTER INTERFACING (BCI) systems have been in the past and still are today motivated by the wish of paralyzed or otherwise disabled people to utilize new means of communication or to extend their mobility. Electroecephalography (EEG) is most often used to capture the electrical brain activity. Examples of these applications are the P300-based spelling systems in the case of communication [1] and the use of imagined movement [2] or slow cortical potentials (SCPs) [3] to control, for example, a wheelchair. The first BCI game was developed by Vidal *et al.* [4]: a simple game in which the user had to navigate through a maze by looking at off-screen fixation points that flashed periodically. However, since then, the available computational power has greatly increased, hardware costs have fallen, and the number of peer-reviewed research articles on BCIs has increased rapidly [5]. From around the year 2000, more and more research groups have developed BCI games. For an overview, see the survey in the article by Plass-Oude Bos *et al.* [6]. As EEG headsets are becoming cheaper [7] and BCI technology more available, the application of BCIs in entertainment is becoming more interesting [8].

A good game provides an immersive experience to users, giving them the feeling they are *in* the game. But BCIs are inherently different from the classical input devices such as a mouse, a keyboard, and a joystick. A BCI provides the user with an unreliable input channel to control the game. Does this hinder immersion and feelings of presence in the game? Can a BCI be of value in a popular modern game? Is it fun to play with BCI control? In this study, we will try to answer these questions by first reviewing related work from previous studies on BCI games and how to evaluate user experience in games through questionnaires and duration estimation. In Section III, we will explain how we incorporated BCI control into the modern game *World of Warcraft (WoW)* and how we designed our experiment.

II. Related Work

A. BCI Games

BCI games have been developed by several research groups, often as proofs of concept or to evaluate the use of mental tasks in an online application. One classic example of a BCI game which also inspired us to some extent is *Brainbal*, developed by Hjelm *et al.* [9], [10]. The concept of the game is to relax more than your competitor. Thus, two players compete to be more relaxed, which makes it a paradoxical and fun game. In this study, the ratio between frontal alpha and beta waves is taken as a measure of relaxation.

Pineda *et al.* [11] developed a BCI for a 3-D shooter game. Forward and backward movements were controlled by the keyboard, turning left and right by alpha levels over the motor cortex. As this study focused on the ability of participants to learn to control their alpha/mu levels over the course of several weeks, the study only included four participants. They found that control over mu activity was easily obtained and maintained.

Instead of focusing only on bit rates and accuracies, Plass-Oude Bos *et al.* applied a user-centered approach and found that the ease of executing a certain mental task is an important factor as well [6]. Gürkök *et al.* found that certain mental tasks for BCI games, while being appropriate, might not be the best option as the major modality for interaction [12].

Manuscript received May 16, 2012; revised November 19, 2012; accepted January 18, 2013. Date of publication March 21, 2013; date of current version June 10, 2013. This work was supported by the BrainGain Smart Mix Programme of the Dutch Ministry of Economic Affairs and the Dutch Ministry of Education, Culture, and Science.

The authors are with the Human Media Interaction Group, Faculty of Electrical Engineering, Mathematics and Computer Science (EEMCS), University of Twente, Enschede P7500AE, The Netherlands (e-mail: laar@cs.utwente.nl).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCIAIG.2013.2253778

B. User Experience

According to Witmer *et al.* [13] both involvement and immersion are necessary for experiencing presence. Witmer *et al.* introduced a presence questionnaire addressing these factors for specific use in virtual environments. Today, games can be at least as realistic as virtual environments. A successful game drags the user along with its immersive graphics, compelling characters, and creative narrative. Users are then able to feel themselves present in the environment or the world that is created by the game.

The sense of presence can be interrupted by distracting factors such as audiovisual stimuli that are not congruent with the virtual world or interfaces that are unnatural or faulty. At the intersection of these games, which aim at providing the ideal circumstances for a feeling of high presence, evaluation of the user experience is key to make games with BCIs successful. Van de Laar *et al.* [14] made an attempt to evaluate a BCI game with a questionnaire based on the game experience questionnaire (GEQ) by IJsselsteijn *et al.* [15] with specific items on the role of the BCI. Van de Laar *et al.* [16] gave an overview of the several methods available to researchers to evaluate BCI games and when to use them. Administering a questionnaire is the best method for quantifying results. To be able to answer the *why* question, interviewing would be a better method.

C. Duration Estimation

When people play a game and are immersed in the experience, they tend to forget the time. This experienced, subjective duration is considered by Waterworth [17] to be influenced by the amount of activity in the working memory. Users who are deeply engaged in experimental tasks or activities such as playing computer games might experience a different duration from those who are not engaged to that extent. Previous studies have not reached a consensus on the direction or the nature of this relation. On the one hand, Waterworth *et al.* [18] suggested that a high level of presence would yield an overestimation of duration. On the other hand, the study by IJsselsteijn *et al.* [19] stated the opposite; in the case of a high level of presence, the user would underestimate the duration; no significant data supported this claim however. Davies [20] suggested that a high level of presence may cause the user to either underestimate or overestimate the duration. In an experiment by Waterworth *et al.* [21] designed to assess the validity of their model posed in [18], participants looked at audiovisual streams with varying content and durations. They found relatively weak relations for some streams between presence and estimated duration. However, the experiment was noninteractive and the longest duration between being interrupted in the experience was 104 s.

III. Methods

A. World of Warcraft

World of Warcraft (WoW) is a massive multiplayer online role-playing game (MMORPG), meaning that thousands of users log on to servers with their virtual characters to battle monsters or each other and evolve their character through an interactive narrative set in a fictional medieval virtual world. The game is immensely popular with currently more than ten million subscribers. The objective of the game is essentially to level up, and get better abilities and better gear. This is done by achieving experience points, gained by completing quests, slaying enemies, and exploring the world.

In our experiment, participants played a character which is part of the Night Elf race and is a druid. The druid has the unique ability to transform itself into bear shape, more or less providing a completely different character.

1) Bear Shape: The bear shape simulates the armor and health similar to a warrior, therefore providing a so-called Tank ability. In this shape, the player is able to soak up damage, especially useful when fighting more than one enemy. At the same time, the bear has slow speed but high impact attacks at close range (melee).

2) Druid Shape: The druid shape is able to cast magic spells on enemies, that is, attacking from a distance. The druid is also able to heal itself and others through spells. More suitable for healing after a fight or when fighting only one enemy from a distance, the druid requires a different playing strategy than when in bear form.

B. α*WoW*

In our adaptation of *WoW*, called α*WoW*, we make use of the power in the alpha band over parietal regions. According to Cantero *et al.* [22], high alpha levels in the parietal lobe indicate a state of relaxed alertness. Also, Barry *et al.* "confirm the arousal link between alpha and electrodermal activity" [23]. We mapped this state of relaxed alertness to the shape of the Night Elf in the druid shape who is strongly dependent on intelligence and mental concentration. The opposite state of relaxed alertness (the decrease in alpha band activity) would be a state of stress or agitation providing a natural relation to the bear which is eager to fight.

The relation of alpha levels to inhibitory processes and passive or idling states of certain parts of the brain is often studied but also often misunderstood, as suggested by Başar *et al.* [24], [25]. According to Başar *et al.* [26] and Pineda *et al.* [11] users can train their alpha levels. This is also what we observed during the many demos we gave of α*WoW* at international (e.g., AISB2009 [27]) and national conferences and popular scientific events (e.g., TEDxAmsterdam 2009 [28]). One person often served as the demo subject and learned to control alpha levels to such an extent that intentionally transforming every 5 s was possible.

C. Experimental Setup

Our experimental setup was mirrored over two rooms because of the large number of participants. Both rooms contained a laptop, including a fast graphics card and a high-definition screen suitable for playing video games. An external mouse was connected to the laptop.

To acquire the EEG data, we used Emotiv's EPOC. The Emotiv EPOC is an inexpensive wireless EEG device. Although inexpensive, it provides more sensors than other inexpensive systems and high usability [29]. It has also been used in other BCI-based games, for example, by Van Vliet *et al.* [30]. We fitted the set on the user's head, tilted at an angle of about 25° (see Fig. 1) to be able to measure the parietal

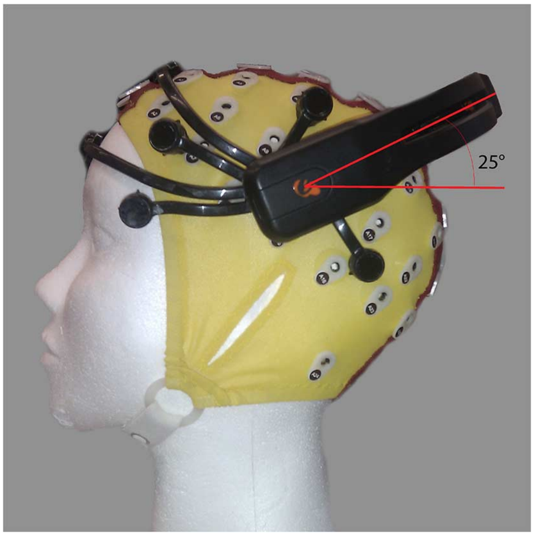


Fig. 1. Photograph of the Emotiv EPOC headset that was tilted forward about 25°.

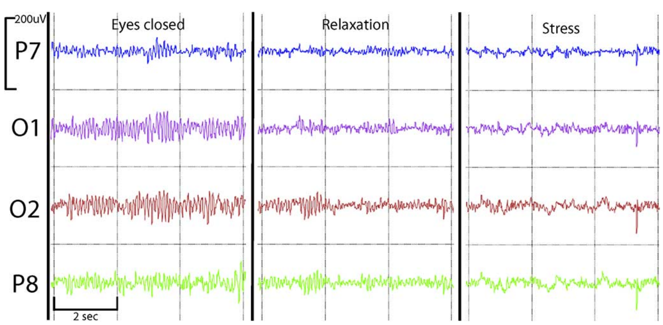


Fig. 2. Examples of EEG data during eyes closed, "relaxation" and "stressed" states.

Fig. 3. Overview of the experimental setup.

Fig. 4. Screenshot of *WoW* with the add-on enabled. The orange bar (top left) provides feedback on the user's current alpha level. Bigger means lower alpha.

brain activity with the O1, O2, P7, and P8 channels. After tilting the headset, these sensors are approximately located at the P1, P2, CP7, and CP8 positions, respectively. Fig. 2 shows randomly taken examples of EEG data segments indicative of a participant in a "stressed" state (high arousal), a "relaxed" state (low arousal), and a segment of data with eyes closed to compare.

A webcam with a built-in microphone was placed in the corners of both rooms to monitor the participant when the experimenter left the room and to record video and audio during the interview at the end of the experiment. An overview of the experimental setup (seen from the webcam) can be seen in Fig. 3.

We used two separate *WoW* accounts on two different servers but within the same game area. We made a default interface layout with offensive and defensive actions recommended for the respective character level. When leveling up a character, it was moved to an area that matched its competence, and possible new actions were added to the action bar. Since the character's armor perishes over time, it was repaired after every session.

1) Control Signal: The pipeline we used to assess the amount of alpha power in the parietal region starts with the acquisition of the EEG data from the Emotiv EPOC headset. The EEG data from channels O1, O2, P7, and P8 are selected and streamed to the analysis pipeline. The online data stream is cut into windows of 2 s with an interwindow interval of 250 ms (4 Hz). For every window, a fast Fourier transform (FFT) is computed for all channels and the band power from 8 to 13 Hz is averaged over all channels. An adaptive Z-score normalization with a time window of 2 min is used to convert the band power into a usable control signal and limit the influence of outliers. A short-term smoothing including the previous two windows is applied with a 20%, 30%, and 50% contribution, respectively.

2) Mapping the Control Signal to Game Actions: The control signal is a 4-Hz signal in the [0, 1] range. To map this signal to actions within *WoW*, we interpret and convert the control signals into virtual key strokes. Every 250 ms the value of the control signal is rounded to the nearest first decimal and sent through a keystroke to the game to supply the user with feedback through means of an animated orange bar, known as the "stress bar" (see Fig. 4). The longer the bar, the lower is the value of the control signal and the lower is the amount of alpha activity. To control the actual shape shifting, we applied hysteresis, that is, direction-specific thresholds of 0.3 and 0.7 to trigger the transformation. As the control signal is normalized, the 0.3 and 0.7 thresholds roughly compare to either one standard deviation below and one standard deviation above the average of the signal. To transform to a bear form, a control signal of 0.3 or lower is needed, to transform from a bear form to a druid form, a control signal of 0.7 or higher is needed.

To make sure the user really intends to transform after 1 s (four windows) of dwell time, an indication of the upcoming transformation is shown. When transforming from a bear form to a druid form, the edges of the screen flash a translucent blue for 1 s; when transforming from a druid form to a bear form, the edges flash a translucent red for 1 s. If the control signal still exceeds the threshold 1 s after the onset of the flash (eight windows in total), the actual transformation is triggered.

D. Participants and Instructions

Forty two participants were recruited for this study. Before the experiment, participants were asked to fill in an online questionnaire indicating whether and if so what experience they had with games and specifically with *WoW*. If they indicated having experience with *WoW*, some additional questions were asked. Twenty eight participants (66.7%) had no previous *WoW* experience, while 14 participants (33.3%) had previous *WoW* experience. The participants' ages ranged from 17 to 49 years (mean: 24.86; SD: 6.88). Twelve participants (29%) were female, and 30 participants (71%) were male. The experimental sessions were done on an individual basis, lasting about 100 min per participant. Participants were paid €12 for their efforts (unless employed by the University of Twente, Enschede, The Netherlands). All participants were asked to sign a written informed consent form prior to the experiment. All instructions were provided in English. The first three authors of this manuscript carried out the experiments.

1) Progression of the Experiment: The participant was welcomed and took a seat at the laptop. To make the time assessment as unbiased as possible, participants were asked to turn off mobile phones and to hand them in, together with their wrist watches, for the duration of the experiment. *WoW* by default has a small on screen clock which we removed through a script at the start of every experiment. Because the desktop of the Windows environment was visible before launching *WoW*, we hid the taskbar. The Emotiv EPOC headset was installed on the participant's head, and the experimenter made sure all sensors were operating at green signal quality levels, which indicates proper impedance levels according to the manufacturer of the headset. The participant received general instructions on playing *WoW* and played a tutorial session of 5 min to get acquainted with the controls and interface. The experimenter answered any questions posed by the participant. After this, either a session with or without BCI control started with the instructions. After the BCI pipeline was started, the normalized control signal was given the time to stabilize for 1 min before the actual session started. If instructions were clear, the experimenter literally told the users to play the game until they did not want to play anymore. Should they wish to stop the current session, the participant was instructed to ring a bell that was on the table he was sitting at. The experimenter left the room and started the stopwatch. If the user did not ring the bell within 30 min, the experimenter interrupted the participant playing the game; if the user did ring the bell, the experimenter immediately noted the time that had passed and went back into the experiment room. In either case, the experimenter immediately administered the questionnaire because the participants had to make an estimation of the duration of the session. After the questionnaire was finished, the experimenter prepared for the second session and gave the participant the relevant instructions. During this time, absolutely no mention of time passed was made, even when asked for by the participant. The second session followed the same course of events as the first. After the second session, the Emotiv EPOC was removed from the participant's head and recording of video and audio started to capture the oral interview that followed.

2) Questionnaire: The questionnaire administered in this study was based on the presence questionnaire by Witmer *et al.* [13]. This questionnaire, although published in 1998, is still the *de facto* questionnaire for measuring presence. According to Van Baren *et al.* [31], most other presence questionnaires are based on the one by Witmer *et al.* and are less well validated. We used the involvement/control scale and the natural scale, as the other subscales in this questionnaire would not provide us with useful information. We also added two scales. The transformation control scale has specific items on the amount of control the user has over the transformation action. The items in this scale are based on [14], specified on BCI-specific issues that would not be captured by the involvement/control scale. The fun/achievement scale was also added to provide us with information on the amount of fun, concentration, and achievement the user experienced.

E. Validation of the Control Signal

To assess the performance users can achieve by means of the BCI pipeline used in the first experiment, we conducted a second experiment in a more controllable and clinical setting. The concept of the game is a ball that continuously falls down and has to be "pushed" into the correct (green) basket; see Fig. 5. The bottom of the screen is divided into two equal parts (baskets) both randomly appearing in either red or green during the fall of one ball. It takes 7 s for the ball to fall down. In this window of 7 s, the participant has to push the ball to the left or to the right, depending on which half of the screen is green. By generating high alpha levels, the ball is pushed to the left; consequently, by generating low alpha levels, the ball is pushed to the right. The same pipeline and control signal as defined in Section III-C is used for this experiment, with the screen divided into ten equal parts, according to ten corresponding levels in the control signal. (0.1, 0.2, 0.3, etc.). Near-real-time feedback is generated at a rate of 4 Hz, as in the first experiment. Whenever the ball ends up in the green part, the score is increased by one. Ten participants were recruited for this second experiment. Eight participants were male, and two participants were female. Participants were not given a monetary reward. Participants trained for three sessions of 20 trials each. After the training, five sessions were done resulting in 100 trials per person.

IV. Results

In this section, we will first assess the internal consistency of the scales we used in our questionnaire. Next, we will review our measure for duration estimation by the participants. Then, we will do an analysis of variance on our data to assess

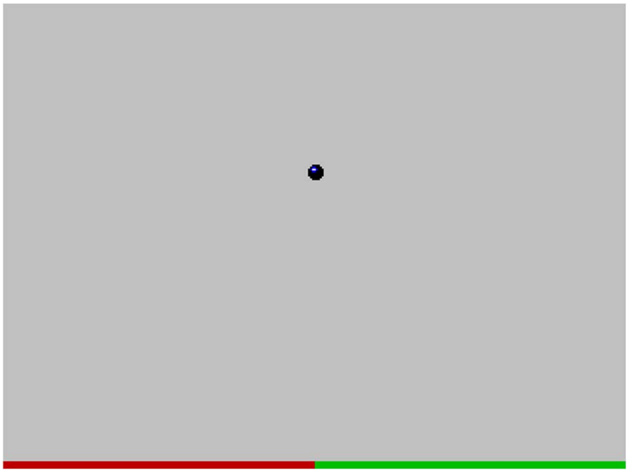


Fig. 5. Screenshot of the ball/basket game.

whether users experienced the condition with BCI control differently (and if so in what way) than the condition without BCI control. Last, we will review the results of the second experiment on how much control is attainable with the BCI pipeline that was used.

A. *Item Analysis*

In our questionnaire, we used the subscales "involvement/control" and "natural" from the presence questionnaire by Witmer *et al.* [13]. The reported Cronbach's alpha for the whole questionnaire was 0.81. In our study, we found the Cronbach's alphas to be a bit lower, 0.733 for involvement/control and 0.655 for natural. We constructed two other scales, one on "transformation control," involving questions on the mechanism to shape-shift specifically, and a scale we would call "fun/achievement," which indicates whether the user had fun, liked to do things he/she could not do in real life, felt a sense of achievement, and was concentrated on the game. Both showed good internal consistency, with a Cronbach's alpha on normalized items of 0.844 and 0.725, respectively.

B. *Duration Estimation*

As explained in Section II-C, the relation between presence and duration estimation by the user is not that clear. Some studies point toward overestimation in the case of high presence, and other studies state that underestimation is more prone to happen. For every participant, an actual (objective) duration and an estimated (by the user) duration were recorded. Comparing the means over all users with a Wilcoxon signed ranks test yielded no significant differences for either the actual duration between conditions ($p = 0.456$) or estimated duration ($p = 0.227$) between conditions. To evaluate whether overestimation or underestimation of duration was more prone to happen, we counted the number of cases in which the difference in durations was bigger than zero and the cases in which the difference was smaller than zero. An overview can be seen in Table I. This table also shows the average amount of overestimation and underestimation.

To be able to see if there were differences between experienced and inexperienced participants, we split up this analysis according to this group factor. An overview of these results can be seen in Table II. This result hints at a slightly better notion of time for experienced participants. The average overestimation is lower and the percentage of participants who estimated the duration correctly is higher for this group.

TABLE I
GRAND AVERAGES OF PARTICIPANTS. DURATIONS PER CONDITION (MINUTES) AND NUMBER OF OVERESTIMATIONS AND UNDERESTIMATIONS OF DURATION

Grand averages		
	BCI	noBCI
avg act	26.1	26.6
avg est	23.5	24.5
over	20	20
exact	7	9
under	15	13
avg over	10.1	11.0
avg under	-6.3	-10.2

TABLE II
AVERAGES OF EXPERIENCED AND INEXPERIENCED PARTICIPANTS. DURATIONS PER CONDITION (MINUTES), NUMBER OF OVERESTIMATIONS AND UNDERESTIMATIONS OF DURATION AND AVERAGE AMOUNT OF OVERESTIMATIONS AND UNDERESTIMATIONS

Experienced			Inexperienced		
	BCI	noBCI		BCI	noBCI
avg act	25.8	26.1	avg act	26.2	26.8
avg est	23.0	23.2	avg est	23.8	25.1
over	8	7	over	12	13
exact	3	5	exact	4	4
under	3	2	under	12	11
avg over	8.0	8.3	avg over	11.6	12.4
avg under	-6.2	-9.5	avg under	-5.8	-10.5

Because there is no clear trend in these results, we took either overestimation or underestimation as a measure using

$$\mathrm{Duration\_diff} = |\mathrm{Duration\_act.} - \mathrm{Duration\_est.}|. \quad (1)$$

Using a Pearson product–moment correlation to correlate this measure to the four scales in the questionnaire, compensating for multiple comparison with a Bonferroni correction yielded no significant results. A Wilcoxon test showed no significant difference in Duration_diff between the two conditions ($z = -1.144$ and $p = 0.259$). However, looking at the Duration_diff measure within subjects over the two conditions showed a strong correlation ($\rho = 0.59$ and $p < 0.001$).

C. *Analysis of Variance*

A 2×2 one-way repeated measures MANOVA was conducted to compare scores on the four scales (involvement, naturalness, transformation control, and fun/achievement) at condition 1 (no BCI) and condition 2 (BCI). Between subject factors are gender (female/male) and experience (yes/no). The overview of this analysis can be seen in Tables III and IV. There was a significant effect for condition (Wilk's lambda

TABLE III
MULTIVARIATE TESTS. EFFECT, WILKS' LAMBDA, F-MEASURE, HYPOTHESIS df, ERROR df AND p-VALUE. "EXPERIENCE" IS BETWEEN SUBJECTS, AND "CONDITION" IS WITHIN SUBJECTS

Multivariate tests					
Effect	Λ	F	df_1	df_2	p
Experience	0.724	3.328	4	35	0.021
Condition	0.548	7.222	4	35	0.001

TABLE IV
WITHIN-SUBJECTS AND BETWEEN-SUBJECTS UNIVARIATE TESTS. SOURCE OF THE EFFECT, EFFECT, MEAN SQUARES, F-MEASURE, HYPOTHESIS df, ERROR df, AND p-VALUE

Univariate tests						
Source	Effect	MS	F	df_1	df_2	p
Condition	Involvement	1.699	11.179	1	38	0.002
	Transf. Control	37.561	21.708	1	38	0.001
Experience	Natural	10.001	5.728	1	38	0.022

TABLE V
SUBJECT ACCURACIES, AVERAGE, STANDARD DEVIATION, AND CONFIDENCE INTERVAL

Per subject performance	
Subject	Accuracy
S1	67%
S2	82%
S3	97%
S4	63%
S5	56%
S6	83%
S7	88%
S8	62%
S9	61%
S10	82%
average	74.1%
stdev	13.9%
CI	9.95%

$= 0.426$, $F(4, 37) = 12.452$, $p < 0.001$, multivariate partial eta squared $= 0.574$). Further investigation of the univariate within-subject tests showed that both the involvement/control scale ($F = 20.369$, $p < 0.001$, partial eta squared $= 0.337$) and transformation control ($F = 40.373$, $p < 0.001$, partial eta squared $= 0.502$) largely caused this effect. No significant interaction effect was found for condition $\times$ experience ($p = 0.421$), condition $\times$ gender ($p = 0.391$) or condition $\times$ experience $\times$ gender ($p = 0.784$). Between subjects, the univariate test with source experience showed a significant effect for the natural scale ($F = 5.655$, $p = 0.022$) with a positive direction, that is, experienced players scored higher on natural. All other tests proved not significant.

D. Second Experiment: Accuracy

All ten participants in the second experiment performed 100 trials. In Table V, we report the accuracy for every subject, the average accuracy (74.1%), the standard deviation, and the corresponding Student's t confidence interval for a sample size of 10.74.1% $+-$9.95% = 64.15% to 78.15%. Random chance level for a binary BCI would be 50%.

V. DISCUSSION

The reliability analysis of the scales showed good internal consistencies for all scales, although somewhat lower than reported by Witmer *et al.* [13]. The cause of this lower internal consistency may be caused by the fact that we used a modern game in our experiment instead of, by today's standards, an old fashioned virtual environment. Another difference is the population of participants that is becoming more and more educated in operating computers, controlling input devices and navigating through 3-D immersive worlds. We constructed the transformation control scale to measure the amount of control over the transformation action. This scale showed a high internal consistency of 0.84 with only six items. A high internal consistency makes sense, considering the specific nature of the scale. It measures the amount of control over the transformation, either with or without BCI. The fun/achievement scale consists of four items on positive effects: having fun, like to do things in the game, being able to concentrate, and having a sense of achievement, concerning concepts broader apart, but still achieving a good internal consistency.

As mentioned before participants could indicate when they wanted to stop playing. The actual duration was recorded as well as their estimation. There was no significant difference in actual duration between conditions, suggesting that participants did not prefer one or the other.

Further analysis of the actual and estimated durations showed no clear overestimation or underestimation of time. This is not in line with what most previous studies reported. Looking at the absolute difference between actual and estimated duration we saw no correlation to any of the scales, supporting the findings by Waterworth *et al.* [21] that duration estimation is not a clear indication of presence. Furthermore, we saw a very strong correlation within subjects. This suggests that duration estimation is largely a personal trait. Further analysis indicated that experienced participants were a little better at estimating the duration. This may well be due to the lower workload involved because of familiarity with the interface and the workings of the game.

The analysis of variance in the form of a one way repeated measures MANOVA showed that in the multivariate case there was a significant difference between playing with BCI control for shape shifting and playing without BCI control (with keyboard/mouse control) for shape shifting. Looking at the univariate test we can see where these differences stem from. The involvement/control scale and the transformation control scale showed a large effect. Both scales were rated lower in the BCI condition, which was to be expected. The amount of perceived control over the transformation was lower in the BCI condition because the control channel was not completely accurate. The user might evoke the right amount of alpha but due to the inherent delay with BCIs (in this case, 2 s), the transformation might happen later than expected and cause the user's avatar to unnecessarily lose hitpoints. In the case of low or no control, the user might also consider the BCI as a hindrance rather than an aid, thus limiting involvement.

Although the amount of perceived control was significantly lower, the fun/achievement scale showed no significant difference between conditions. Our data suggest (through the MANOVA analysis) that fun in this case is not deteriorated by the lower amount of control in the BCI condition. This may have several reasons.

1) Participants found the EEG headset and BCI a novel and fun gadget to play with, notwithstanding the fact that it might have provided them with less involvement in the game. This might be due to the novelty effect of using brain activity and EEG devices in general. Some participants also stated something along these lines during the interview. Either they wanted to see if they could get better at controlling the transformation process, or they said they liked participating in the experiment but would not pay to play with a similar setup, including the BCI, at home. This possible novelty effect is something that is hard to prove or disprove. It can cloud user experience evaluation of BCI games in general. The only way to investigate whether, and if so to what extent, the novelty effect is influencing these results is to perform several follow-up experiments with the same participants.
2) For the users for which it worked better, the BCI provided an added modality to aid them playing the game. Several experienced *WoW* players stated in their interview that they could see such BCI as a nice challenge when playing with a low-level character. They also felt that their amount of control would improve over time, providing them with a useful extra modality in a game which always has a shortage of buttons to express the plethora of functionalities in such a complex game.
3) For the participants who experienced no or little control, the BCI did not interrupt their experience to such an extent that it became a frustration. Some participants also reported this. "I just kept on playing and tried to quickly adapt. At a certain point I memorized all the shape-specific functions and their respective buttons."

There was a between-subjects effect of experience on the naturalness of the game. This result was to be expected. Experienced players are already familiar with the workings of the game, the controls for navigation, the style of interaction with the virtual world the game is set in, and perhaps even the type and race of the virtual character used in the experiment. One experienced participant remarked something about this during the interview. The question (which is in the natural scale) about how natural was the way of moving through the virtual environment was somewhat unclear. Of course using arrow keys is not natural, but if one does this every day, it becomes second nature.

The second experiment in which we evaluated the performance of the BCI pipeline in a more clinical setting showed that on average participants achieved an accuracy of 74.1%. This is significantly higher than one would expect from a system operating on random choices. Looking into the per-subject accuracies, there were two distinct groups: one group (five participants) who got barely above the chance level control (56%–67%) and one group (five participants) which achieved good to excellent control (82%–97%). As the participants are not the same as the participants who participated in the first experiment, we cannot analyze the effect of a low or high accuracy on user experience. However, this result shows that control with the current BCI pipeline is definitely possible, but whether the user has good control is subject dependent.

This study was largely based on one game, and the results we found might not generalize to other kinds of games. This is one of the problems within the field of human–computer interaction (HCI) research in general. In prototype development, certain design choices are made. The type of game, interaction style, aesthetical choices, etc., are all choices that can have an influence on the outcome of the evaluation. As the number of choices and the resulting number of combinations of choices are far bigger than the possibilities for evaluation of these choices, it is hard to generalize or even be conclusive about any result. However, we think that the results we presented in this paper point in certain directions that might be useful in the field of BCI-controlled games.

VI. Conclusion

In this study, we set out to answer the question as to whether BCI control can bring added value to a modern game. We recruited experienced *WoW* players and participants without any experience with *WoW*. We compared the actual duration of the two conditions (with and without BCI control), and our data showed they equally liked to play with and without BCI control. The questionnaires provided us with some more insights. Control and involvement was lower in the BCI condition, but fun was not significantly different in the two conditions. Several reasons might explain this: participants found the BCI a novel and interesting modality to play with, it provided them with a challenge, and/or they could cope with or anticipate the lack of control. Whether this novelty effect wears off and users lose interest remains to be seen. Even without perfect control the addition of BCI control could make a game that gets users curious and interested in this new modality.

Acknowledgment

The authors would like to thank L. Packwood for proofreading this paper.

References

[1] L. Farwell and E. Donchin, "Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalogr. Clin. Neurophysiol.*, vol. 70, no. 6, pp. 510–523, 1988.
[2] R. Leeb, D. Friedman, G. Müller-Putz, R. Scherer, M. Slater, and G. Pfurtscheller, "Self-paced (asynchronous) BCI control of a wheelchair in virtual environments: A case study with a tetraplegic," *Comput. Intell. Neurosci.*, 2007, Article ID 79642.

[3] N. Birbaumer, T. Elbert, A. Canavan, and B. Rockstroh, "Slow potentials of the cerebral cortex and behavior," *Physiol. Rev.*, vol. 70, no. 1, pp. 1–41, 1990.

[4] J. Vidal, "Real-time detection of brain events in EEG," *Proc. IEEE*, vol. 65, no. 5, pp. 633–641, May 1977.

[5] J. Wolpaw and E. Wolpaw, "Brain-computer interfaces: Something new under the sun," *Brain-Computer Interfaces: Principles Practice*, p. 3, 2012, DOI: 10.1093/acprof:oso/9780195388855.001.0001.

[6] D. Plass-Oude Bos, B. Reuderink, B. L. A. van de Laar, H. Gürkök, C. Mühl, M. Poel, A. Nijholt, and D. K. J. Heylen, "Brain-computer interfacing and games," in *Brain-Computer Interfaces. Applying our Minds to Human-Computer Interaction*, ser. Human-Computer Interaction, D. Tan and A. Nijholt, Eds. London, U.K.: Springer-Verlag, Jul. 2010, pp. 149–178.

[7] B. Zhang, J. Wang, and T. Fuhlbrigge, "A review of the commercial brain-computer interface technology from perspective of industrial robotics," in *Proc. IEEE Int. Conf. Autom. Logistics*, 2010, pp. 379–384.

[8] Future BNCI, "A roadmap for future directions in brain/neuronal computer interactions research," 2012 [Online]. Available: http://future-bnci.org/images/stories/Future_BNCI_Roadmap.pdf

[9] S. Hjelm and C. Browall, "Brainball-using brain activity for cool competition," in *Proc. NordiCHI*, 2000, pp. 177–188.

[10] S. Hjelm, "Research+design: The making of brainball," *Interactions*, vol. 10, no. 1, pp. 26–34, 2003.

[11] J. Pineda, D. Silverman, A. Vankov, and J. Hestenes, "Learning to control brain rhythms: Making a brain-computer interface possible," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 181–184, Jun. 2003.

[12] H. Gürkök, G. Hakvoort, and M. Poel, "Evaluating user experience in a selection based brain-computer interface game: A comparative study," in *Proceedings of the 10th International Conference on Entertainment Computing (ICEC 2011)*, ser. Lecture Notes in Computer Science, J. Coutinho Anacleto, S. Fels, N. Graham, B. Kapralos, M. Saif El-Nasr, and K. Stanley, Eds. Berlin, Germany: Springer-Verlag, Oct. 2011, vol. 6972, pp. 77–88.

[13] B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: A presence questionnaire," *Presence: Teleoper. Virtual Environ.*, vol. 7, no. 3, pp. 225–240, Jun. 1998.

[14] B. L. A. van de Laar, B. Reuderink, D. Plass-Oude Bos, and D. K. J. Heylen, "Evaluating user experience of actual and imagined movement in BCI gaming," *Int. J. Gaming Comput. Mediated Simul.*, vol. 2, no. 4, pp. 33–47, Oct. 2010.

[15] W. IJsselsteijn, Y. de Kort, K. Poels, A. Jurgelionis, and F. Bellotti, "Characterising and measuring user experiences in digital games," in *Proc. Int. Conf. Adv. Comput. Entertain. Technol.*, 2007, pp. 1–4.

[16] B. L. A. van de Laar, H. Gürkök, D. Plass-Oude Bos, F. Nijboer, and A. Nijholt, "Perspectives on user experience evaluation of brain-computer interfaces," in *Universal Access in Human-Computer Interaction. Users Diversity*, ser. Lecture Notes in Computer Science, C. Stephanidis, Ed. Berlin, Germany: Springer-Verlag, Jun. 2011, vol. 6766, pp. 600–609.

[17] J. Waterworth, "Memory mechanisms and the psychophysical scaling of duration," *Perception*, vol. 14, no. 1, pp. 81–92, 1985.

[18] E. Waterworth and J. Waterworth, "Focus, locus, and sensus: The three dimensions of virtual experience," *Cyber Psychol. Behav.*, vol. 4, no. 2, pp. 203–213, 2001.

[19] W. IJsselsteijn, I. Bierhoff, and Y. Slangen-de Kort, "Duration estimation and presence," in *Proc. Presence Conf.*, 2001, DOI: 10.1.1.19.5785.

[20] C. Davies, "Landscape, earth, body, being, space, and time in the immersive virtual environments osmose and ephémère," in *Women, Art, and Technology*, J. Malloy, Ed. Boston, MA, USA: MIT Press, 2003, pp. 322–337.

[21] J. A. Waterworth and E. L. Waterworth, "Being and time: Judged presence and duration as a function of media form," *Presence: Teleoper. Virtual Environ.*, vol. 12, no. 5, pp. 495–511, Oct. 2003.

[22] J. Cantero, M. Atienza, C. Gómez, and R. Salas, "Spectral structure and brain mapping of human alpha activities in different arousal states," *Neuropsychobiology*, vol. 39, no. 2, pp. 110–116, 1999.

[23] R. J. Barry, A. R. larke, S. J. Johnstone, C. A. Magee, and J. A. Rushby, "EEG differences between eyes-closed and eyes-open resting conditions," *Clin. Neurophysiol.*, vol. 118, no. 12, pp. 2765–2773, 2007.

[24] E. Başar, M. Schürmann, C. Başar-Eroglu, and S. Karakaş, "Alpha oscillations in brain functioning: An integrative theory," *Int. J. Psychophysiol.*, vol. 26, no. 13, pp. 5–29, 1997.

[25] E. Başar, C. Başar-Eroglu, S. Karakaş, and M. Schürmann, "Gamma, alpha, delta, and theta oscillations govern cognitive processes," *Int. J. Psychophysiol.*, vol. 39, no. 23, pp. 241–248, 2001.

[26] E. Basar, C. Basar-Eroglu, R. Parnefjord, E. Rahn, and M. Schürmann, "Evoked potentials: Ensembles of brain induced rhythmicities in the alpha, theta and gamma ranges," in *Induced Rhythms in the Brain*. Boston, MA, USA: Birkhäuser, 1992, pp. 155–181 [Online]. Available: http://dx.doi.org/10.1007/978-1-4757-1281-0_9

[27] Demonstration of Alpha WoW at AISB 2009, Edinburgh, Scotland.

[28] Demonstration of Alpha WoW at TEDxAmsterdam 2009 [Online]. Available: http://www.tedxamsterdam.com/2009/alphaworldofwarcraft/

[29] K. Stamps and Y. Hamam, "Towards inexpensive bci control for wheelchair navigation in the enabled environment a hardware survey," in *Brain Informatics*, ser. Lecture Notes in Computer Science, Y. Yao, R. Sun, T. Poggio, J. Liu, N. Zhong, and J. Huang, Eds. Berlin, Germany: Springer-Verlag, 2010, vol. 6334, pp. 336–345 [Online]. Available: http://dx.doi.org/10.1007/978-3-642-15314-3_32

[30] M. van Vliet, A. Robben, N. Chumerin, N. Manyakov, A. Combaz, and M. Van Hulle, "Designing a brain-computer interface controlled videogame using consumer grade EEG hardware," in *Proc. IEEE Biosignals Biorobot. Conf.*, 2012, DOI: 10.1109/BRC.2012.6222186.

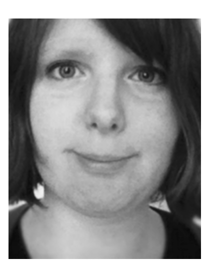
[31] J. van Baren and W. IJsselsteijn, "Measuring presence: A guide to current measurement approaches," *OmniPres Project, Deliverable 5* 2004 [Online]. Available: http://www8.informatik.umu.se/~jwworth/PresenceMeasurement.pdf

Bram van de Laar received the B.Sc. degree in computer science and the M.Sc. degree in human media interaction from the University of Twente, Enschede, The Netherlands, in 2007 and 2009 respectively, where he is currently working toward the Ph.D. degree in brain–computer interfaces and user experience.

He was a Visiting Researcher at the MIT Media Lab, Cambridge, MA, USA, with the Information Ecology group in 2012. His research interests include brain–computer interfaces, multimodal interaction, and user experience design and evaluation.

Hayrettin Gürkök received the B.Sc. and M.Sc. degrees in computer engineering from Bilkent University, Ankara, Turkey, in 2006 and 2008, respectively, and the Ph.D. degree from the University of Twente, Enschede, The Netherlands, with a dissertation on user experience evaluation and brain–computer interface games.

His research interests include entertainment computing, user experience, and hedonic brain–computer interfaces.

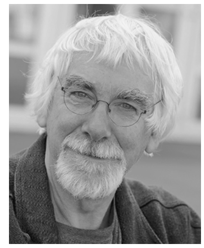


Danny Plass-Oude Bos received the B.Sc. degree in computer science and the M.Sc. degree with a thesis on the user experience of using BCI for games from the University of Twente, Enschede, The Netherlands, in 2006 and 2008, respectively. Currently, she is working toward the Ph.D. degree in the Human Media Interaction Group, University of Twente.

She gained experience in brain–computer interfaces by looking at artifact detection in an online EEG system at the Radboud University, Nijmegen, The Netherlands.

Mannes Poel received the Ph.D. degree in mathematics from the University of Utrecht, Utrecht, The Netherlands, in 1986.

After that he was a NWO Research Fellow at the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, for one year, after which he became a Postdoctoral Researcher at the University of Groningen, Groningen, The Netherlands. In 1989, he was appointed as an Assistant Professor in the Faculty of Computer Science, University of Twente, Enschede, The Netherlands. His current research topics involve the application of machine learning to vision-based human behavior detection and brain–computer interfacing.



Anton Nijholt studied mathematics and computer science at the Technical University of Delft, Delft, The Netherlands and received the Ph.D. degree in theoretical computer science from the Vrije Universiteit, Amsterdam, The Netherlands, in 1980.

He is a Professor of Computer Science in the Human Media Interaction Group, University of Twente, Enschede, The Netherlands. He held positions at various universities in and outside The Netherlands. His main research interests are multimodal interaction, affective computing, brain–computer interfacing, and entertainment computing.

Prof. Nijholt has been General and Program Chair of the main international conferences on affective computing, virtual agents, and entertainment computing.