Actual and Imagined Movement in BCI Gaming

Bram van de Laar

March 27, 2009

Master's Thesis
Human Media Interaction
EEMCS Faculty
University of Twente

Graduation Committee:
Christian Mühl (daily supervisor)
Danny Oude Bos
Boris Reuderink
Dirk Heylen
Mannes Poel

Abstract

Abstract

Most research on Brain-Computer Interfaces (BCI) focuses on developing ways of expression for disabled people who are not able to communicate through other means. Recently it has been shown that BCI can also be used in games to give users a richer experience and new ways to interact with a computer or game console. This thesis describes the research conducted to find out what the differences are in user experience between actual and imagined movement as paradigms in a BCI game. A group of twenty subjects played the game BrainBasher controlling their actions in the game purely with brain waves. Questionnaires were administered throughout the whole experiment to measure user experience. BrainBasher also yielded a statistic on how well people were able to control the game. The questionnaire was corrected for internal consistency and used to evaluate the differences. Results showed that users found more challenge in imagined movement but at the same time stayed more alert when performing actual movement. Furthermore the brain activity related with actual movement was found to be better recognizable by the classification pipeline used by BrainBasher.

Contents

1

Introduction

1.1 Motivation

Brain-Computer Interfaces (BCI) have been a topic of research for quite some time now. The focus of research is recently shifting from quality of life improvement for handicapped people to the entertainment of healthy people. At the Human Media Interaction (HMI) department at the University of Twente this trend was anticipated by collaborating with other universities, research institutes and companies in the BrainGain project. At HMI we focus on the interaction of humans with a BCI in relation to games. The original goal in previous BCI research was to provide handicapped users to communicate through a BCI with only their minds. While this is a very useful approach, healthy people might also find joy in using a BCI through actual motion of their limbs. This approach seemed underexposed in previous research. While some research has been conducted on recognizing actual movement and there are some comparisons of the clinical, neurological differences between actual and imagined movement [MMVW00]

When looking at the success of the Nintendo Wii, it becomes clear that actual movement is already well enjoyed by gamers [Hea08]. Moreover, imagined movement in adulthood is not as trivial as actual movement is. Although for example professional sportsmen and musicians use imaginary movement for training an actual motor skill it still is not as trivial to do as actual movement [Jea94]. Though one can think of various applications in which imagined movement is used, these are almost always associated with skills which require a lot of training. Actual movement might therefore be a more natural and easier way of interacting with a BCI.

In the field of BCI, brain activity is recorded and automatically interpreted to be applied in various applications. Measuring brain activity is already well known in medicine using the electroencephalographam (EEG). EEG is a proven method, which has a number of advantages over other methods: it is non-invasive, has a high temporal resolution, does not require a laboratory setting, is relatively cheap, and it is even possible to create wireless EEG head-sets. BCI systems need to make decisions based on very short segments of EEG data to make it useful for different applications such as wheelchairs, robots, and personal computers. In the case of software applications, BCI can be used as an additional modality of control, for evaluation of the user or the application, or to build adaptive user interfaces [NTA+08].

Games are usually the first applications to adopt new paradigms, driven by the gamers' continuing search for novelty and challenges [NTP+08]. Apart from them being a suitable platform to bring this new interaction modality to the general population, games also provide a safe and motivational environment for patients during training or rehabilitation [GAG07, LLK+07]. Research has shown that using BCI instead of the conventional mouse and keyboard can add to the user experience by making a game more challenging, richer, and more immersive [BR08].

Before BCI can be adopted by the general population there are still a number of issues that need to be addressed: artifacts in the recorded brain data (signals that do not stem from the brain), inter and intra-subject variability, inter and intra-session variability, long training periods, low transfer rates (of commands), and BCI illiteracy [SBM08]. Apart from that, more attention from the Human Computer Interaction community is required on how this new input modality influences the user experience, and how the interaction can be improved [LLR+08]. While most research into using movement for BCI has focused on imagined movement for usage in Brain Computer Interfaces, some clinical research shows that actual movement in fact elicitates a more pronounced and therefore better discernable signal in the motor cortex [MMVW00].

Most conducted research in BCI is characterized by the following aspects: The focus is mainly on users with a handicap such as an evolved state of ALS, e.g. people that cannot communicate through other means. This study however will focus on healthy users. Most research has used experiments which were conducted in clinical environments and in which the sole means of the study is to evaluate the BCI *an sich*. Mostly the brain activity data is only gathered and processed at a later time. This is the complete opposite of this study, in which users will receive direct feedback during the experiment and not the BCI is studied, but how the users experiences it. The direct feedback guarantees that the interaction for the user will be a lot richer. The last aspect in which this study is different from most other BCI research is that the aim was to evaluate a large pool of subject for the experiments. While some studies only report on three to ten subjects that were well recognizable by their BCI system, this study uses twenty participants and also reports on all of them.

1.2 Introduction to BCI

Brain-Computer Interfaces are literally meant as a means of communication between the brains of a human being and a computer. The human brain consists of roughly 10^{11} neurons which are all really small cells that can generate electrical impulses used to signal other neurons. Because each neuron can connect with thousands of other neurons, the human brain is probably the most complex electrical network in the world. While one might expect that the result of all these tiny electrical impulses yields to chaos, the neurons tend to fire within groups in a wave like pattern, generating the so called brain waves. These brain waves, which are the addition of thousands of electrical impulses, can be measured as small potentials by various techniques. These brain waves of course do not contain all the information that is passing through the neurons, but are a kind of echo of what happens inside the brain. Brain-Computer Interfaces are specially targeted at distilling useful information from these signals and exploiting it to control computers with.

Input Methods

Nowadays different techniques exist to extract and visualize the signals generated by the human brain. Every technique has its own pro's and cons as will be explained in the following paragraphs.

fMRI funtional Magnetic Resonance Imaging (fMRI) is, in contrast to what was explained in the previous section not a technique that measures the electrical currents generated by the neurons, but measures the blood flow passing through a particular area in the brain. This amount blood flow indicates where the most activity in the brain is at that particular moment. The sites in the brain that are activated can be linked to different functional regions (see for example the motor cortex and its functions in Figure 1.1). The advantage is that an fMRI has a fairly high spatial resolution ($\pm$1mm) but the disadvantage is a low temporal resolution ($\pm$0,5Hz). Furthermore the equipment is quite expensive (millions of euro's) and not portable.

MEG Magnetoencephalography (MEG) does not measure the potentials generated by the neurons directly rather the magnetic fields induced by them. This way a fair spatial resolution ($\pm$2mm) and a good temporal resolution ($\pm$1KHz) is achieved. A big disadvantage for usage in a BCI is that a magnetically shielded room is needed for measurements. Furthermore the equipment is like the fMRI equipment rather expensive and not portable.

EEG Electroencephalography (EEG) measures the electrical potentials present at the scalp of a person. This is achieved by sticking highly sensitive electrodes on the subjects scalp, while using electrolyte gel assuring a good electrical connection. While the spatial resolution depends on the number of sensors and is fairly low ($\pm$5mm) a very high temporal resolution can be achieved (up to 4 KHz and higher). EEG equipment can be very affordable, especially when mass produced and can be portable. This makes it an ideal candidate to be used in a BCI. A disadvantage that exists is that setting up an EEG has its practical issues such as making the gel connections and connecting the sensors, which can take some time.

ECoG Electrocorticography (ECoG) can be seen as 'the next step' after EEG. In essence it uses the same techniques as EEG, only now directly on the brain cortices (by means of a transplant) instead of on the scalp. This way the spatial resolution will be boosted. On the downside a high risk operation is needed, medicines need to be taken to suppress rejection of the transplant and the transplant is almost always targeted at a small part of the brain.

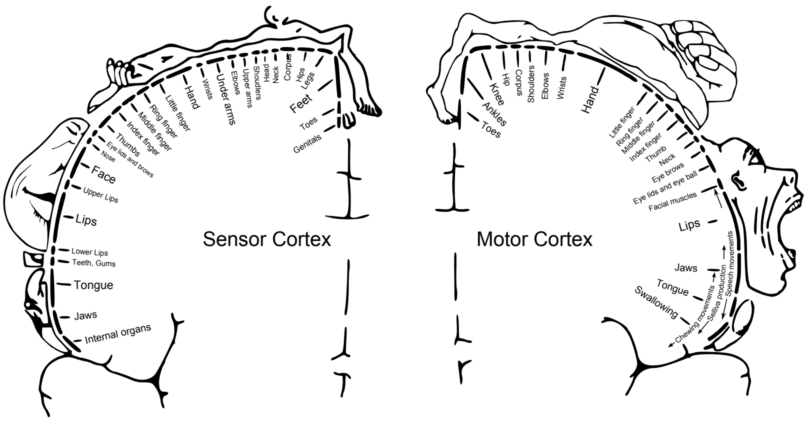

Figure 1.1: The Homunculus, depicting functional brain regions.

Signals and characteristics

The signals that can be measured by EEG are various but always have wave like properties. The most important ones will be explained below.

P300 The P300 signal is a positive deflection and a event related potential (ERP). The signal can be measured from about 300ms after the onset of a stimulus which triggers the attention of a person. It can be used in various applications, the most widespread probably the P300-speller.

ERS/ERD Event Related (De)Synchronization is the change of band power. Desynchronization of the signal lowers the band power, synchronization increases the band power. ERS and ERD are often related to activity in the (pre-)motor cortex. Using the Homunculus (see Figure 1.1) one can determine what motor cortex activity caused the ERD or ERS. An important aspect of ERD/ERS is that it is not only seen when there is actual motor activity but also when someone imagines to move a hand for example. This is called motor imagery or imagined movement.

Bereitschaftspotential The Bereitschaftspotential of Readiness Potential (RP) [KD65] is a signal in the brain activity that leads up to voluntary movement of a person. It is associated with the planning of the movement and happens in a point in time (50ms) before the actual movement is carried out.

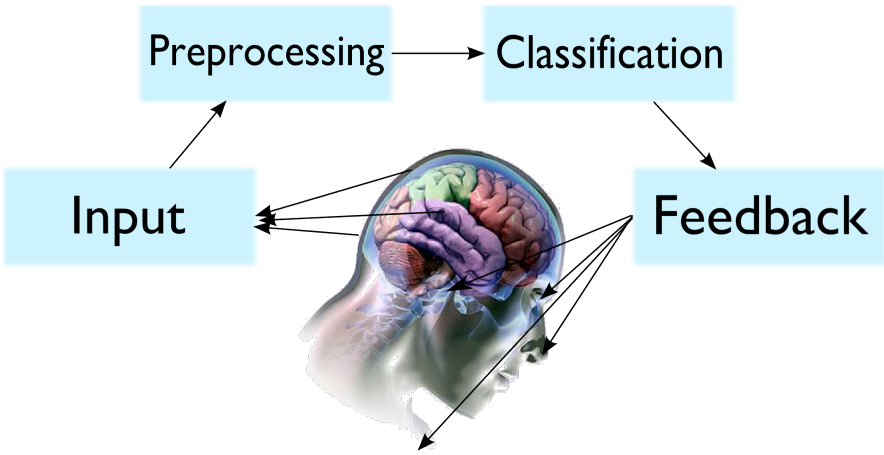

Figure 1.2: Overview of a BCI system.

Although the signal is quite weak, when mastered, this signal could be used for a immensely fast BCI that recognizes what you meant, before you even moved. [KDB+04]

BCI in practice

A BCI in practice will consist of a multiple components. First of all, the brains of the person that wants to use the BCI, the equipment to measures the brain activity, the apparatus that receives the signal and converts it to a digital form and the computers that preprocess and analyse the signals. The last step is to provide the user with feedback of what the BCI did. This whole cycle can be seen in Figure 1.2. Every step in this cycle is important and constitutes to a BCI system. In terms of research at the HMI group we focus not only on the classification but also on the way feedback is presented to and experienced by the user.

Why Imagined and Actual movement

Previously a lot of research has been done in aid of handicapped people, exploring new methods to aid these people in expressing themselves and improving their quality of life. Often these groups of handicapped people partially or completely 'locked-in' as a sympton of ALS, Parkinsson or MS. With EEG equipment becoming more cost effective and processing power of computer increasing, one can also look at BCI's for a broader target audience, especially for entertainment purposes such as video games, neuro-feedback applications and serious games.

A few BCI games based on imagined or actual movement do already exist. Pineda et al. designed a first-person shooter game in which the user could move using the keyboard, and turn by imagined movement [PSVH03]. Players learned to control the BCI by experimenting; no instructions were given beforehand. Other examples include the virtual environments of Leeb et al. [LKF+05], the board game of Kayagil et al. [KBL+07], and the game BrainBasher [BR08] that was used in this study.

Sometimes BCI research is dogmatic and stuck with the idea that imagined movement is the only way to go. While this is entirely true for handicapped patients, healthy users might find actual movement easier, more encouraging etc. Furthermore, a phenomenon called BCI illiteracy is apparent in at least 1 out of 6 persons [SBM08] and [BPK+]. BCI illiteracy means that a subject is not capable to interact through a BCI at all. Because most studies use imagined movement, actual movement might prove to be a more reliable signal for a BCI to use.

Both actual and imagined movement can be used for BCIs. Apart from these two kinds of movement there exists another kind of movement, called quasi-movement. Quasi movement was first presented bij Nikulin et al [NHJC08]. It is defined by the kind of movement that is volitional but not detectable by other objective means as for example a motion tracking system. Obviously, actual movement is the most natural and intuitive way for users to communicate and express themselves. All the games which involve movement tasks are based on a neurological process known as Event-Related Desynchronization (ERD)[Pfu01]. ERD is detectable as a decrease in power in the β-frequency band on corresponding motor cortices. Before use the BCI has to be adapted to person-specific examples of the ERD using machine learning techniques.

1.3 User experience

An important part in this study will be the user experience during the BCI game. User experience of games is still a somewhat underexposed topic in current research. There are some attempts at defining the concepts that are important in user experience by Larsen [Lar08] and Desurvire et al. [DW08]. The former looks at how game reviewers have reviewed games in the last twenty years in a quantitative as well as in a qualitative approach. The latter carries out a real evaluation, but the intended purpose for this evaluation is making games more approachable and not evaluating the game *an sich*. IJsselsteijn et al. developed the Game Experience Questionnaire [IdKP+07]. This questionnaire is explicitly aimed at evaluating the users experience, directly asking the users how they feel afterwards and what they felt during the game. This questionnaire consists of several scales which cover most concepts defined by Larsen. While this questionnaire is a step in the right direction for evaluating a BCI game, some concepts that are of key importance for a BCI game are missing from this questionnaire. On the other hand, some concepts are not applicable to BCI games because there is no story line or apparent flow in the current BCI games. Concepts that are in the GEQ are: Positive Affect, Negative Affect, Flow, Sensory Immersion, Tension, Challenge, Competence and Tiredness.

1.4 Research Questions

Building upon the previous paragraphs, the following research questions can be formulated.

The most important question in this thesis is: "What are the differences in user experience between actual and imagined movement in a BCI gaming environment?" An important factor that probably constitutes to a difference in user experience is that actual movement will probably yield a better performance of the BCI system. Because experiments are carried out in an interactive setting, a better performance of the BCI system will have a direct influence on the user's experience.

While research has been conducted on the clinical differences between actual and imagined movement in brain activity [MMVW00], no clear results are available for an actual application of these two different modalities. The hypothesis is that actual movement is better recognizable through a more pronounced ERD in the beta band in the motor cortex. One could say that imagined movement in the brain looks like an attenuated version of actual movement. This would possibly lead to a better discernable signal when one uses actual movement in a BCI.

Looking at a broader target audience one could ask if everyone is able to use a BCI? Are there restrictions in age, does it only work for students and is it also possible for young children and elder people to use a BCI? Are there any differences between men and women? Anecdotal findings suggest that women would achieve better recognition rates than men. This probably is not the case. There is no scientific evidence that women would produce a better recognizable signal in their motor cortex through ERD/ERS. This thesis focusses on a balanced proportion of men and women in the subject pool to evaluate a possible difference.

2

Methods

The main question in this thesis is whether there are differences between imagined and actual movement in a BCI gaming environment. To this extent an experiment was carried out with twenty subjects. The use of materials such as the BCI setup and the questionnaire that was used are covered in this chapter. Also the methods that were used tot analyse the data will be explained in this chapter.

2.1 Materials

The materials used for conducting the experiment consist of five important parts: The BCI setup, containing the EEG equipment and the computers used for playing the game and for analysis. The BCI setup was used to interact with the BCI game BrainBasher. Particpants had to sign an informed consent, fill in a demographical survey and their user experience was evaluated through a User Experience Questionnaire. There will be an elaboration on each part below.

BCI setup

EEG equipment The EEQ equipment that is available at the HMI department and which was used for the experiment consists of a BioSemi ActiveTwo 32 channel AD box with 32 active Ag/AgCl sensors. The sensors are active in the sense that a high carrier signal is used to modulate the signal on. This way noise is kept to a minimum.

Setup The EEG equipment is connected to the recording pc for data acquisition through an USB port. The marker box is connected to the presentation pc for capturing the markers it sends out. See Figure 2.1 The EEG data together with the markers are recorded on the recording PC and at the same time sent to the classification server for analysis. The classification server provides the game on the presentation PC with classifcation results based on the EEG data which are feed back to the user in an graphical way.

PC's The PC that was used during the experiment for recording purposes consisted of a Intel Pentium 4 3.2GHz CPU, 1GB of RAM. The presentation PC consisted of an Intel Pentium 4 2.8GHz CPU, 1GB of RAM and a 19 inch monitor. The EEG data was recorded for future reference, but the data was also sent through to a computation server to be analysed online. This computation server consisted of two quadcore Intel Xeon's running at 2.8GHz and had 24GB of RAM.

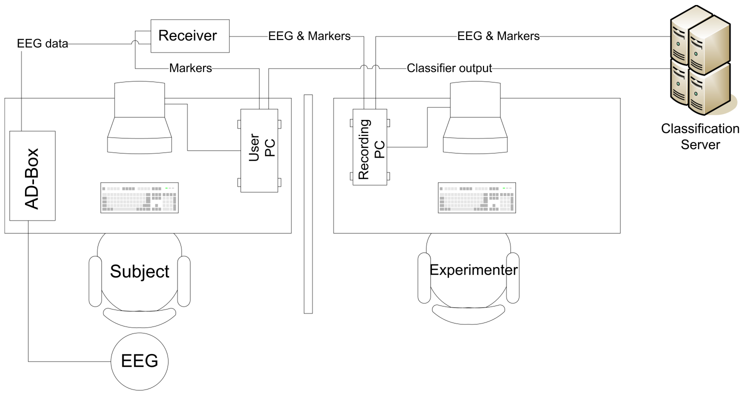

Figure 2.1: Overview of the BCI setup

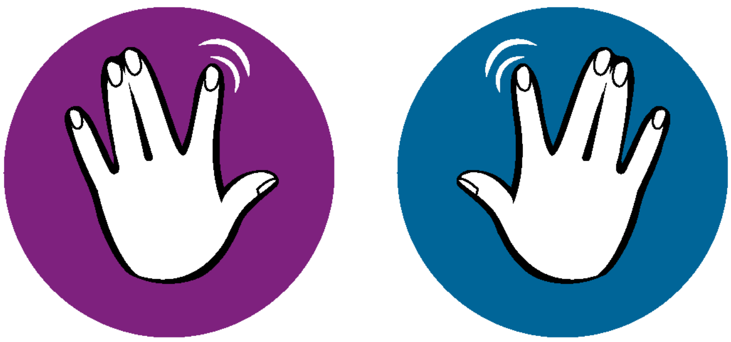

Figure 2.2: The symbols for left and right hand movement. (Courtesy of Danny Oude Bos.)

BrainBasher

The BCI game used for this research is *BrainBasher*[BR08]. The goal of this game is to perform specific brain actions as quickly as possible. For each correct and detected action you score a point. Game control is achieved by two mental tasks: left hand movement versus right hand movement. For the actual movement task both hands are laid on the desk in front of the user. When the appropriate stimulus appears they have to perform a simple tapping movement with their whole hand. When performing the imagined movement task users are instructed to imagine the same movement, without actually using any of their hand muscles.

Before the user can play however, they will have to undergo a training in which stimuli (in the form of symbols denoting the user actions, see Figure 2.2) and breaks are alternated. During the stimulus the subject performs the indicated action: movement of the left or right hand. The user is instructed to stay relaxed and not to move, excluding the break periods, to prevent noise

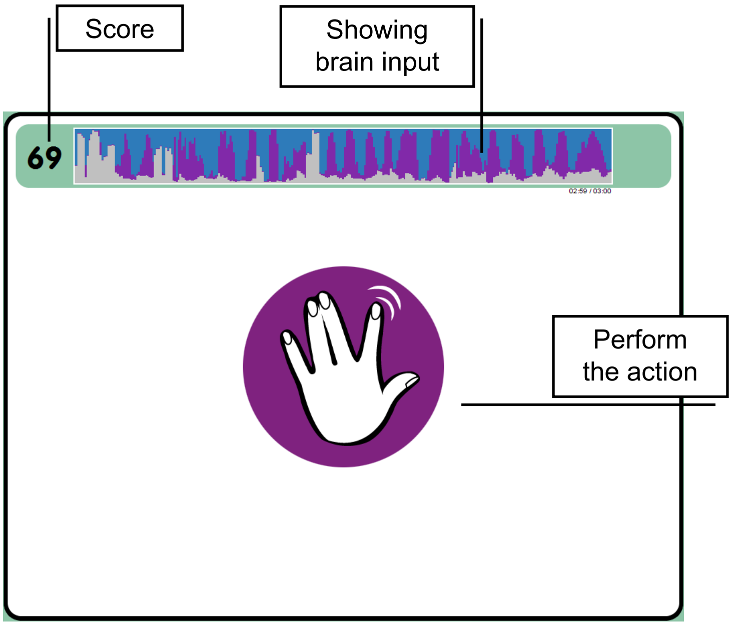

Figure 2.3: A game session. (Courtesy of Danny Oude Bos.)

in the EEG. This is of course with the exception of the hand movement in the case of the actual movement sessions. In our system, the training consists of two short sessions, taking ten minutes in total. The EEG data from both training sessions are concatenated and used for training the classifier of the BCI system.

During the game session the user is instructed to take care that they carry out precisely the same movement (be it actual or imagined) as in the preceding training session. The difference is that they have to react as fast as they can to each new symbol popping up by performing the action right away. Bashing a symbol is accomplished when the classifier recognizes the action, according to a confidence level of at least 60%. Every *bash* results in one point added to their total score. The goal of the game is to bash as many symbols in the allocated three minutes, to achieve a maximum score (Figure 2.3.)

A schematic view of the total system is shown in Figure 2.4. The user interacts with the system by executing brain actions, and also by keyboard to traverse the menu. Brain activity is acquired with a BioSemi EEG setup using 32 electrodes, at a 256Hz sample frequency. For training the system, examples of the ERD for both the left hand and right hand are used to derive a linear classifier to be used during the online game session. The EEG is processed as follows. First the raw data is re-referenced to the common average (CAR) to remove external sources of noise. After re-referencing a bandpass-filter isolates the frequency range in which the ERD occurs. Then we train spatial

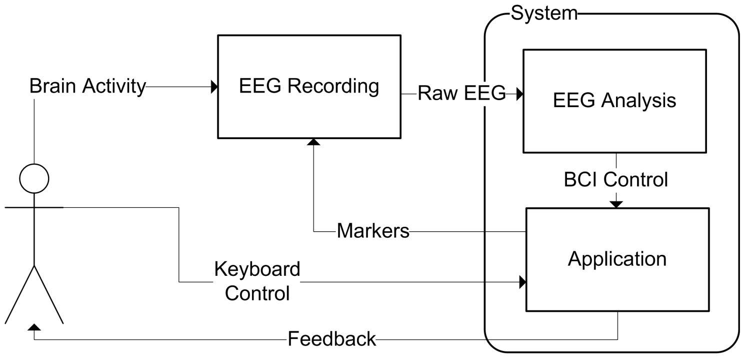

Figure 2.4: BrainBasher System View. (Courtesy of Danny Oude Bos.)

filters with the Common Spatial Patterns (CSP) algorithm [Kol91] to the motor cortices. These spatial filters are used to extract the band power in the most discriminative sources. Linear Discriminant Analysis (LDA) is applied to make a final prediction based on the band power features. After training the BCI generates four new predictions every second, based on the real-time EEG data. These predictions are used to control the game.

Informed Consent

An informed consent is: "[..] a legal condition whereby a person can be said to have given consent based upon a clear appreciation and understanding of the facts, implications and future consequences of an action." [?] This means that the participant has to read and sign a consent form stating everything the user needs to know to take part in the experiment, his or her rights and what happens with the acquired data. Therefor, the first thing that happened after the participants were seated is that they read and signed the informed consent. All signed informed consents are copied, ordered and archived for future reference. The exact informed consent as administered during the experiment can be found in Appendix A.

Demographical Survey

The demographical survey consists of general demographics similar to other fields of research, as well as some specific questions on caffeine and tobacco consumption habits. Furthermore various items on the user's alertness and fatigue are incorporated. For the demographical survey as it was administered during the experiment see Appendix B.

Questionnaire

To evaluate the user experience a questionnaire based on the GEQ [IdKP+07] is developed. The original GEQ consists of the scales Alertness, Challenge, Competence, Flow, Sensory Immersion, Negative Affect, Positive Affect and Tension as can be seen in Table 2.1. Although the GEQ consists of a lot of useful questions for evaluating various games, its main purpose is evaluating complex and immersive 3D virtual games. Therefore the questionnaire has been adapted to evaluate the user experience especially in BCI games.

For most scales only the name of the scale was altered or one or two items were left out to keep the questionnaire as short as possible. Items that tended to be applicable to longer, more complex or taking place in a virtual world were left out. These were prone to produce identical results for both kinds of movement.

Furthermore there were two scales that were not applicable to BCI games, at least to the one used for this study. The scales Flow and Tension were completely left out of the adapted questionnaire. Flow is defined by a set of elements but boils down to be characterized by degree of control over the game and the sense of being *into* the game. While this study provided the user with a realtime feedback on their actions, there always was an inherent delay and the BCI never was (even near) a 100% recognition rate. Therefore the Flow scale was replaced by the Control scale, which involves better defined items for BCI games. The Tension scale in the original GEQ measures the overall tension and level of irritation that is evoked by the game. However, BrainBasher is a fairly simple and short game without a lot of features. Most irritation will arise from the BCI component which is assessed by the added Control scale (see next paragraph). Hence this scale was not used.

On the other hand some concepts exist that are not applicable for normal games but that do play an important role when playing a BCI game. The different way of interacting with the game, through a BCI, gives the game itself a different dimension. Because a BCI is not a failsafe device such as a combination of keyboard and mouse, the amount of control a user experiences is an important factor. When a user is unable to achieve a high recognition percentage via the BCI this obviously has its effects on their positive affection towards the system. The Control scales was added to evaluate the sense of control a user experiences. Using a BCI is more strenuous than using a keyboard, it costs a fair bit of concentration to use a BCI. Staying concentrated is tiring. Therefore the Fatigue scale was added. In combination with the Alertness scale this will give a good indication of the users attention and concentration.

The difference between the original GEQ and the adapted BCI version can be seen in Table 2.1

The actual questionnaire consists of statements to which users can respond on a 5 point Likert-scale ranging from 'completely disagree' to 'completely

Comparison of original and adapted GEQ			
Original Scale	No. of items	Adapted Scale	No. of items
-	-	Alertness	4
Challenge	5	Challenge	5
-	-	Control	4
Competence	5	Goals	5
Tiredness	2	Fatigue	4
Flow	5	-	-
Sensory Immersion	6	Immersion	7
Negative Affect	5	Negative Experiences	6
Positive Affect	5	Positive Experiences	7
Tension	5	-	-

Table 2.1: Orginal GEQ scales (left) and adapted GEQ scales (right).

agree'. Some examples of items in the questionnaire are:

1. "I liked playing the game."
2. "I felt the computer recognized my actions."
3. "I'm exhausted."

The complete questionnaire with all items ordered to scale can be seen in Table 2.2. This questionnaire is in Dutch for obvious reasons. One participant with German as a mother tongue attended the experiment. The questionnaire was therefore translated to German and translated back to Dutch to check for and correct inconsistencies. For the administration of the questionnaire during the experiments all items were ordered in a randomized fashion to avoid a possible tendency of participants to fill in the questionnaire on 'auto-pilot'. The final questionnaire exactly as it was administered during the experiments can be seen in Appendix C.

Table 2.2: All items and scales.

Item Code	**Item**
Negatieve ervaringen (Nx)	
N1	Ik vond het tijdverspilling.
	Continued on next page

Table 2.2 – continued from previous page

Item Code	Item
N2	Ik had het gevoel dat ik betere dingen had kunnen doen.
N3	Ik heb spijt van het spelen.
N4	Ik voel me er slecht over.
N5	Ik raakte gerriteerd.
N6	Ik vond het saai.
Positieve ervaringen (Px)	
P1	Het voltooien van de taak voelde als een overwinning.
P2	Ik voelde me trots.
P3	Ik voelde me tevreden.
P4	Ik voelde me machtig.
P5	Ik voelde me voldaan.
P6	Ik vond het leuk om het spel te spelen.
P7	Ik voelde me energieker.
Moeheid (Mx)	
M1	Ik voel me vermoeid.
M2	Ik ben uitgeput.
M3	Ik kreeg hoofdpijn tijdens het spelen.
M4	Ik kon me goed concentreren.
Uitdaging (Ux)	
U1	Ik voelde me uitgedaagd.
U2	Ik vond het moeilijk.
U3	Ik voelde me onder druk gezet.
U4	Ik was gefrustreerd.
U5	Ik moest er veel moeite in steken.
Immersion (Ix)	
I1	Ik vond het moeilijk om terug te keren naar de werkelijkheid.
I2	Ik voelde me gedesorienteerd.
I3	Ik had het gevoel dat ik terugkwam van een reis.
I4	Ik was mijn gevoel voor tijd kwijt.
I5	Ik vond het een nieuwe ervaring.
I6	Ik vond het indrukwekkend.
I7	Het design van het spel sprak me niet aan.

Continued on next page

Table 2.2 – concluded from previous page

Item Code	Item
Controle (Cx)	
C1	Ik had het gevoel dat ik het spel bestuurde.
C2	Ik had het gevoel dat het spel mij bestuurde.
C3	Ik had het gevoel dat de computer mijn bewegingen herkende.
C4	Ik had het gevoel dat de computer me feilloos aanvoelde.
Doelen (Dx)	
D1	Het doel van het spel is duidelijk.
D2	Het spel heeft een steile leercurve.
D3	Ik begreep wat ik moest doen.
D4	Goed scoren werd gedurende het spel steeds makkelijker.
D5	Ik had snel door hoe ik het spel moest spelen.
Alertheid (Ax)	
A1	Ik besefte hoe goed ik presteerde tijdens het spel.
A2	Ik kon me goed focussen op de acties die ik moest doen.
A3	Ik was met andere zaken bezig.
A4	Ik moest moeite doen om me te concentreren.

2.2 Experiment Setup and Procedure

An experiment has been carried out in which users communicate with the BCI game BrainBasher[BR08] using both kinds of movement. First, users fill in a form regarding demographics including handedness as well as characteristics that could influence their ability to focus on the task (like alcohol and caffeine consumption habits). This data is used to check for group differences during analysis. Our experiment consists of two parts: Actual movement and imagined movement. The order of performing actual and imagined movement is randomly assigned for each subject. Each part consists of two sessions.

For the system to be able to recognize the user's actions, a training session is required to create a user-specific classifier. For this classifier a cross-validated error rate is calculated which gives an indication of how well the system is able to detect the actions from a specific user. This is followed by a game session, after which the subject fills in a user experience questionnaire. This questionnaire has been designed based on the GEQ developed at the Game Experience Lab in Eindhoven [IdKP+07]. With this information the user experience for actual and imagined movement can be compared. Between all sessions are breaks in which the user can relax for a minute or two. An overview of the sequence of events during the experiments can be seen in Figure 2.5.

The setup is situated in a normal office environment in contrast to a shielded room. This setting was chosen deliberately as it is a more representative setting for home use. Besides, the EEG system used has active

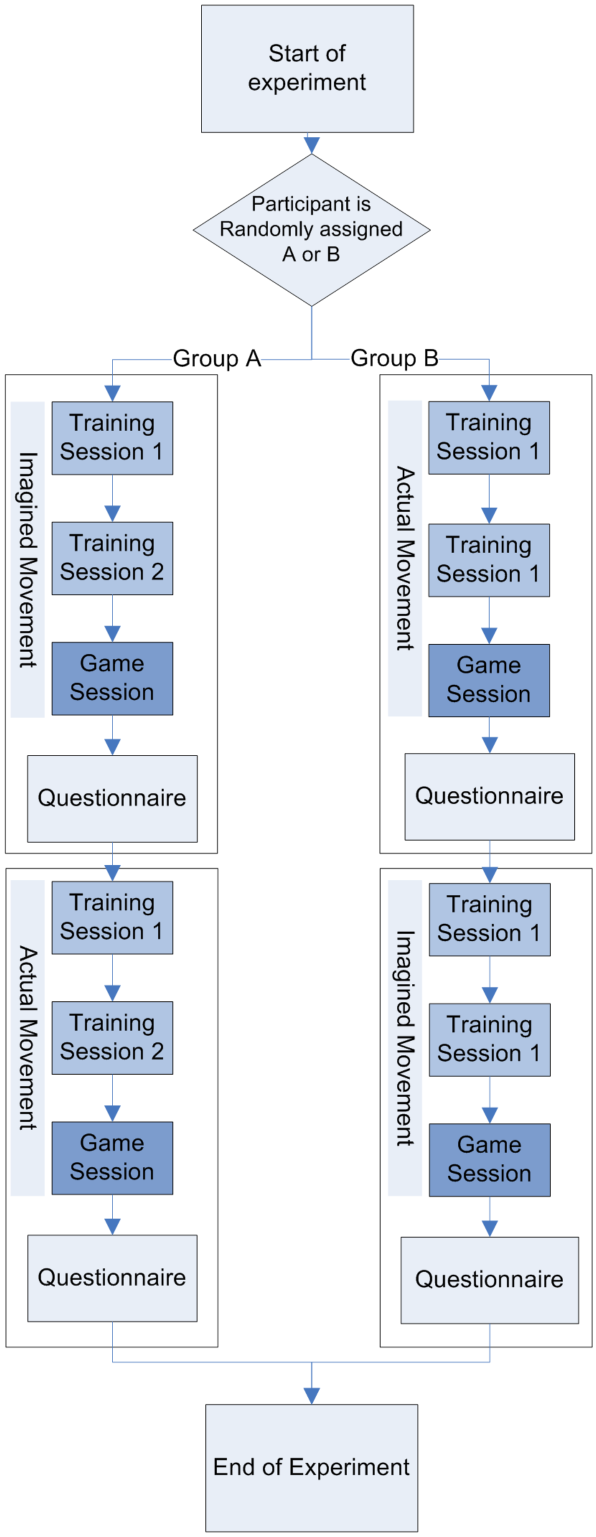

Figure 2.5: Overview of the sequence of events in the experiment.

electrodes which pick up a lot less noise than passive electrodes would. During the experiments themselves, only the researcher will stay in the room. This way distractions will be kept to a minimum, while still being able to provide help when needed.

The experiment is set up as a randomized cross-over experiment to eliminate sequence and learning effects induced by the succession of both tasks. The adapted questionnaire will also be evaluated and possibly corrected so it can be used for assessments of other BCI games and modalities. After all experiments are done the results of the actual movement sessions versus the results of the imagined movements game sessions are compared.

Protocol

The outcome of a BCI experiment is subject to a lot of variables such as user concentration and fatigue, interference from electromagnetical devices, distraction by external sounds, movements and other sensory excitations, hence there can be a high inter-session variability. It is therefore of paramount importance that as much of the controllable variables are kept exactly the same during the course of the experiments. To eliminate as much unwanted variability in the data as possible, an protocol for the experiment was constructed and followed during all experiment sessions. The protocol is split up in a User part and a System part. The User part refers to all actions happening with and around the user, the system part accompanies all other actions the experimenter has to carry out. The complete protocol can be seen in Appendix D.

2.3 Participants

An important part of the BCI setup is of course the participant that is using the BCI. Because this study used two groups in a randomized crossover experiment, at least twenty participants are needed for statistical analyses to be meaningful. This way, each group consists of ten persons. The aim was to get as much men as women into the subject pool. This way it is possible to check for a statistical difference between men and women. The total group consisted of twenty participants, ten male and ten female. Five of them were left handed and 15 were right handed.

2.4 Questionnaire Construction

In order to verify the hypotheses on the research questions in this thesis, the data that is acquired through means of the questionnaire has to be analysed with a number of statistical tests. The questionnaire has to be checked and possibly reconstructed for reliability, the resulting scales will also be checked for homogeneity through factor analyses. Differences between the scales can be found by conducting non parametric paired tests.

Scale construction

To analyze if the different scales that were originally in the GEQ are reliable in the setting, a reliability analysis will be carried out. This analysis will consist of different components: To analyze the results of the questionnaire, we will use Cronbach's Alpha [Cro51]. Alpha is a measure of internal consistency. It is (to a certain extent) a measure of how reliably a scale constructed out of the selected items will measure one concept. This does not necessarily mean that you are measuring the concept you intended to measure, therefore further (qualitative) validation is needed. Alpha is only an estimator of reliability: it measures to what extent the different items are correlating and are consistent, taking subject and environment variance into account. In this research the Standardized Alpha will be used because we want to sum standard scores to construct scales from Likert scale items. A commonly accepted threshold for alpha of 0.7 is the goal for every scale, with 0.6 as a bare minimum [Cor93].

2.5 Data Analysis

Factor Analysis

Factor analysis is a widely used technique with various applications. Factor analysis is mostly used as an Exploratory Factor Analysis (EFA) with Principal Components Analysis (PCA) to reduce the data to a certain number of underlying factors. In other words: the variance in the data can be explained through a limited number of Principal Components (PC's). The amount of variance explained depends on the Eigenvalues of the correspondence matrices. Furthermore, an orthogonal rotation, like the Varimax rotation, maximizes the variance of the squared loadings of a factor on a few variables in a factor matrix. This way it is easier to identify each set of variables with a different factor **??**. Another, less frequently used, method would be to conduct a Confirmatory Factor Analysis (CFA) to see if the presupposed modes in the GEQ are useful in this setting according to the numerical data. This is, in a way, a form of Structured Equation Modelling (SEM). Then a common factor analysis (cFA) would be more appropriate to see if the prior knowledge about the models loads onto the same factor as was hypothesized a priori. This method yields different factor loadings and one can then use (heuristically chosen) cut-off points for the loading factors to see what fits the data best.

Scree Plots

Scree plots can be used for an heuristical evaluation of the number of relevant components in a factor analysis. These plots are constructed by plotting the value of each successive eigenvalue against the rank order. The point at which a strong 'knee' occurs can be used as a cut off point for useful components. The knee is the part where the difference of the derivative of the line (if the connected points were a function) is the largest. This also often is the point where the eigenvalue drops below 1. [CV77]

Analysis of questionnaire data

To analyse which items are significantly different between actual and imagined movement the student t-test will be used. One could also test if there is a significant learning effect by taking one group (e.g. group A) and comparing the results to a second test where the other group (group B) is the population. Skewness and Kurtosis tests can point out if the distribution is Gaussian. This is necessary for the t-test because it assumes a Gaussian model.

Analysis of performance data

The Cross Validation error on the training set is determined by the Brain-Basher classifier. This statistic indicates to what extent the classifier of the BrainBasher system was able to correctly recognize the user's actions based on the training data. Because this is evaluated through a n-fold cross validation it is a good estimation of the performance of the participant in BrainBasher. The statistic consists of three percentages: left vs right, right vs rest and left vs rest. Only the first one is of real importance because it depicts the ability to differentiate between left and right hand movement. An error percentage of 50 resembles the same performance as if the classifier would simply guess independent of the presented data, because of the equal class distribution. An error percentage of 0 would mean a perfect BCI system. With data from all participants of their performance in BrainBasher it is possible to analyze if there is a significant difference in recognizability between imaginary and actual movement through a non-parametric paired test such as Wilcoxon's signed rank paired test. This is a non-parametric statistical hypothesis test for related samples or in this application for two measurements on a single sample. One can see the two error rates as two observations of the same sample (subject). One observation is after Task A (e.g. imagined movement) and the other observation is after Task B (e.g. actual movement). Because the experiment was set up as a randomized crossover experiment, sequence and learning effects (if there are any) will be averaged out.

3 Results

3.1 Descriptive statistics

First we describe the demographics of the test subject pool, then we analyze the questionnaire used for evaluation. Using the results from the questionnaire we can look at the differences in user experience between actual and imagined movement.

Participants

Twenty healthy persons participated as test subjects in this study. The average age across the group was 26.8 (standard deviation: 12.3, minimum: 13 and maximum: 58). Of the twenty participants ten (50%) were male and ten (50%) were female. 19 (95%) participants were Dutch, 1 (5%) participant German. Apart from standard demographics we also asked participants their handedness, because this characteristic might be of influence: 5 (25%) participants are dominantly left-handed, 15 (75%) are right-handed. 14 (70%) received a high education. Computer usage and game experience varied a lot among participants: 8 (40%) participants use a PC for more than six hours a day, 5 (25%) use a computer on a less than daily basis. The same variance goes for game experience: 2 (10%) game two hours a day, 8 (40%) on a weekly basis, 6 (30%) on a monthly basis and 4 (20%) never played a game.

Test subjects were randomly assigned to either group A or group B. Group A would do imagined movement as their first task and actual movement second, group B would do exactly the opposite. Each group had ten (50%) participants.

3.2 Analyses of user experience questionnaire

Questionnaire analysis

All participants filled in the questionnaire after both tasks without missing any questions. The responses on the same items for both tasks were stored in the same respective variables for scale analyses and in separate variables to analyze the differences in user experience between both tasks. Scale analysis was carried out in order to evaluate if the newly developed questionnaire would be useful as a reliable tool to assess user experience in BCI games. The total user experience questionnaire consisted of 42 items over 8 scales. Each item consists of responses to a statement on the user's experience on a 5 point Likert-scale.

Some items were recoded to avoid an expected negative correlation. This was done to questions that were asked in a contra indicative way. An example of such a question is item A3 in the Alertness scale: "I was busy doing other things." A concurring answer to this question is contra indicative of being

Construction of Scales		
	No. of items	Standardized Alpha
Alertness	4	0.575
Challenge	5	0.777
Control	4	0.502
Goals	5	0.583
Fatigue	4	0.673
Immersion	7	0.504
Negative Experiences	6	0.569
Positive Experiences	7	0.865

Table 3.1: Scales before corrections including Standardized Alpha

alert and concentrated. Therefore the values of these questions are reversed, e.g. '1' becomes '5' and '5' becomes '1'.

A first reliability analysis of every scale yielded the results as can be seen in Table 3.2.

Analysis per scale

When looking at the values of the Standardized Alpha some scales obviously need attention. Every scale is discussed in detail in the following paragraphs. The criteria for deletion of an item for the correction of a scale will be an α of 0.6 preferably 0.7 and an item-total correlation >0.3. However, these are just guidelines; for example a item-total correlation $< .3$ will give support for deleting an item from a scale but this will be considered per scale in the next paragraphs.

Alertness The Alertness scale (see Table 3.2) consists of items on the level of alertness and concentration the user is experiencing. The original scale had an α of 0.575 and consists of four items of which two were recoded. Item A1 exhibits clear inconsistent behaviour compared to the scale as a whole. With a negative item-total correlation it is the first candidate for elimination. The Cronbach's Alpha is deleted indicates that α will be around 0.780 which is excellent. The Alertness scale now consists of three items.

Challenge The Challenge scale already exhibits exemplary consistency with an α of 0.777. One might argue about deleting 'U1' because it has a item-total correlation < 0.3 and also will load low on the first component in a factor analysis. However, the item 'U1' is the most important question in the scale

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
A1	10,88	10,522	-,045	,014	,780
A2	10,00	7,179	,594	,550	,295
rec(IA3)	9,58	8,558	,387	,307	,461
rec(IA4)	10,50	5,077	,615	,636	,175

Table 3.2: Alertness: Item Total Statistics

(see Table 2.2) and for the purposes of this study an even higher α of 0.85 is not needed, so this items remains in the scale. The scale therefore remains unchanged consisting of five items.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
U1	9,82	20,199	,169	,115	,852
U2	10,38	13,010	,746	,752	,671
U3	11,58	17,379	,573	,430	,745
U4	10,80	13,805	,731	,602	,680
U5	10,02	15,102	,638	,690	,716

Table 3.3: Challenge: Item Total Statistics

Control The Control scale was specifically added to this questionnaire to measure the amount of control the user experiences when interacting with a computer through a BCI. "To what extent do I control the system and does it understand me?" The scale has the lowest initial Alpha but also an item that has the lowest item-total correlation. While this is a recoded item, which was meant to measure contra indicatively, one might consider to not recode this item for higher correlation with the other items. This however still does not yield a very well fitting item. Either way this item showed a low or negative item-total correlation. Therefore the item is deleted from the scale resulting in an α of 0.783.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
C1	8,73	4,769	,529	,338	,150
rec(IC2)	6,90	8,862	-,181	,094	,767
C3	8,23	4,794	,379	,455	,302
C4	8,92	5,302	,598	,507	,152

Table 3.4: Control: Item Total Statistics

Goals The Goals scale measures to what extent a user understands the game and knows what tot do. The scale has an α of 0.583 with at first sight one item which prevents the scale from achieving a higher Alpha. Deleting item 'D2' improves α to 0.65, which is average. Because 0.7 was the goal for every scale and there are still four items in the scale, a further look into the new item-total statistics revealed that deleting another item would improve the internal consistency greatly. While in Table 3.5 'D5' might seem the most appropriate candidate, because of the lower item-total correlation, 'D4' is the item that is deleted. The squared multiple correlation was already low and with deleting item 'D2', item 'D4' became 'detached' from the scale. Deleting these two items results in an α of 0.754.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
D1	12,60	8,656	,485	,527	,430
D2	14,35	11,003	,019	,105	,654
D3	12,65	7,823	,601	,626	,356
D4	14,60	8,195	,326	,131	,504
D5	13,30	8,062	,292	,311	,532

Table 3.5: Challenge: Item Total Statistics

Fatigue The Fatigue scale intends to measure the amount of fatigue or exhaustion the user is experiencing. An initial α of 0.673 is not bad, but deleting item 'M3', which has a low item-total correlation improves α to 0.759.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
M1	5,67	4,328	,697	,500	,434
M2	6,12	7,189	,521	,349	,604
M3	6,55	8,408	,220	,111	,745
rec(IM4)	5,72	5,897	,503	,319	,595

Table 3.6: Fatigue: Item Total Statistics

Immersion The Immersion scale is an attempt to measure to what extent users feel immersed into the gaming experience. Are they really *into* the game? Beforehand this looked like a difficult scale to construct, especially for BCI games, which games are generally not that immersive. (e.g. no photo-realistic graphics, no music etc.) This proved to be a valid hypothesis. The initial Immersion scale exhibits low internal consistency and the rightmost column does not show a straight way out of this. The heuristic of deleting all items with an item-total correlation <0.3 while maintaining at least three items in the scale provided the solution. The corrected Immersion scale consists of 3 items, 'I4', 'I5' and 'I6' with an α of 0.620. While not perfect, this is the best attempt at measuring immersion with the given questionnaire and data set.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
I1	16,70	13,292	,050	,072	,573
I2	16,50	11,949	,289	,139	,508
I3	16,53	12,512	,150	,185	,549
I4	15,58	8,507	,466	,262	,408
I5	13,90	10,554	,320	,179	,489
I6	14,98	8,794	,501	,326	,393
rec(I7)	13,68	12,533	,107	,142	,566

Table 3.7: Immersion: Item Total Statistics

Negative Experiences The Negative Experiences scale seems a trivial one, measuring to what extent users dislike the game and are not happy playing it. Deleting item 'N2' improves the α from 0.569 tot 0.638. As we will also see in the next section (3.2.2.1), the negative experiences are a bit difficult to grasp with this questionnaire. Probably this is because there can be numerous underlying reasons why someone dislikes the game. Because there is no other solution, only 'N2' is deleted from the scale and the α below 0.7 is taken granted.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
N1	9,90	9,990	,438	,545	,472
N2	9,25	11,321	,084	,099	,626
N3	10,45	12,459	,146	,209	,575
N4	10,12	10,010	,361	,264	,499
N5	9,20	7,908	,444	,286	,446
N6	9,70	9,138	,417	,459	,467

Table 3.8: Negative Experiences: Item Total Statistics

Positive Experiences The Positive Experiences scale measures the positive affection the user experiences, e.g. did they like playing the game? With an α of 0.865 this scale exhibits exemplary behaviour. Deleting an item would only lower the internal consistency of the scale.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
P1	16,90	29,323	,563	,360	,854
P2	16,98	27,717	,809	,739	,816
P3	16,30	26,933	,687	,654	,835
P4	17,65	32,644	,603	,496	,849
P5	16,80	28,574	,696	,623	,833
P6	15,72	32,410	,543	,324	,854
P7	16,65	30,695	,568	,490	,851

Table 3.9: Positive Experiences: Item Total Statistics

Remarks on Cronbach's Alpha

The reader might have noticed by now a discrepancy between the "Cronbach's Alpha if Item Deleted" column and the reported Alpha when this item actually is deleted. This is because the reported Alpha is the Standardized Alpha which deviates somewhat in value from Cronbach's Alpha. The Standardized Alpha is used because the items in the scales are summed in combination with the used Likert-scale. [Cor93].

Resulting scales

Correcting the scales for items that did not constitute to the scales consistency, e.g. deleting items with a low or negative Inter-Item Correlation, Cronbach's Alpha's ranged from 0.620 tot 0.865 and all scales still consisted of at least three items (see Table 3.2). The resulting scales can be seen in Tables 3.10 through 3.17.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
A2	7,22	5,615	,663	,549	,668
rec(IA3)	6,80	6,626	,498	,300	,822
rec(IA4)	7,72	3,333	,782	,636	,525

Table 3.10: Alertness: Item Total Statistics after corrections

Factor analyses

To evaluate the usefulness and dimensionality of the resulting scales, a factor analysis was done on all scales separately. The factor analysis tries to 'pull' the items that constitute to the variance of the scale in certain dimension out of each other. The result is a table with factor loadings on the extracted components for every item. The scree plots are representations of the factor

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
U1	9,82	20,199	,169	,115	,852
U2	10,38	13,010	,746	,752	,671
U3	11,58	17,379	,573	,430	,745
U4	10,80	13,805	,731	,602	,680
U5	10,02	15,102	,638	,690	,716

Table 3.11: Challenge: Item Total Statistics after corrections

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
C1	4,70	4,626	,539	,326	,754
C3	4,20	3,754	,598	,418	,710
C4	4,90	4,810	,711	,507	,606

Table 3.12: Control: Item Total Statistics after corrections

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
D1	7,80	4,472	,533	,525	,661
D3	7,85	3,618	,741	,622	,416
D5	8,50	3,385	,435	,284	,834

Table 3.13: Goals: Item Total Statistics after corrections

analyses and plot the eigenvalues of the resulting factor matrix versus component number. The first dimensions (equivalent to the first component) in the factor analyses of every scale explained more than 55.8% of the variance in the data except for the Negative experiences scale (see Table 3.2).

Scree plots

Scree plots [CV77] indicated strong unidimensionality across all scales except for the Negative Experiences scale, which turned out to be a two dimensional scale (see Figure 3.2. An example of a unidimensional scale can be seen in Figure 3.1.

The corrected questionnaire consisted of 32 items divided over 8 scales. An overview of all corrected scales with their respective Cronbach's Alpha's and variance explained by the first dimension in factor analysis can be found in Table 3.2. The variance explained by the first factor measures to what extent a scale is measuring only one underlying attribute or construct.

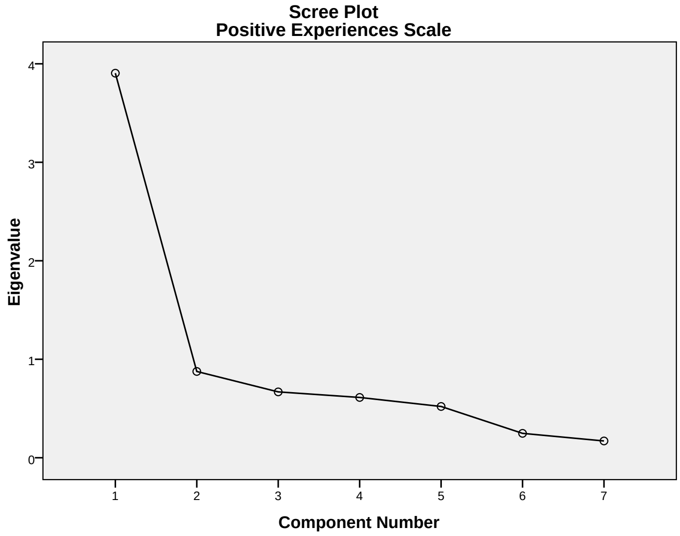

Figure 3.1: Example of a unidimensional scale. (Positive Experiences)

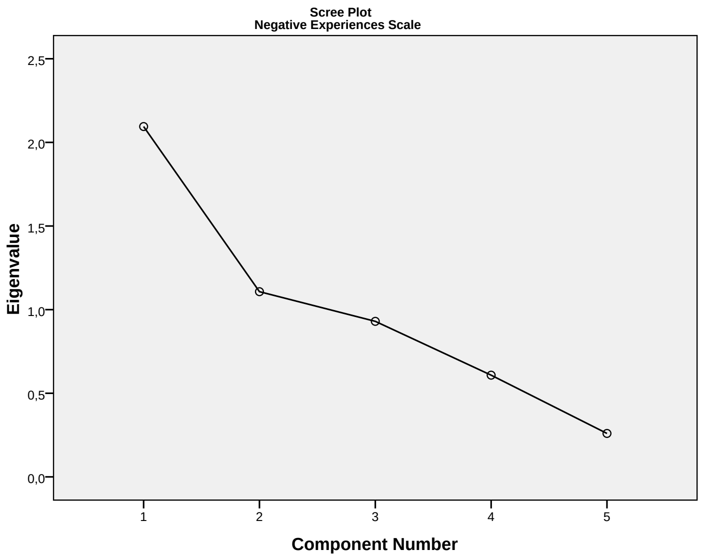

Figure 3.2: Example of a multidimensional scale. (Negative Experiences)

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
M1	4,20	3,138	,661	,446	,562
M2	4,65	5,362	,567	,343	,704
rec(IM4)	4,25	4,141	,556	,315	,679

Table 3.14: Fatigue: Item Total Statistics after corrections

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
I4	7,07	3,866	,425	,198	,531
I5	5,40	4,964	,368	,146	,599
I6	6,48	3,948	,502	,252	,410

Table 3.15: Immersion: Item Total Statistics after corrections

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
N1	7,42	7,635	,548	,538	,498
N3	7,98	10,128	,219	,202	,637
N4	7,65	8,028	,380	,264	,572
N5	6,72	6,615	,372	,253	,598
N6	7,22	7,204	,441	,455	,539

Table 3.16: Negative Experiences: Item Total Statistics after corrections

3.3 Differences in user experience and performance

The final corrected scales were used to compare the user experience for users performing both kinds of movements. The descriptives of all scales can be seen in Table 3.18 A direct comparison by means of paired t-tests was done. The results of these tests can be seen in Table 3.2. The first column is the difference of the means of both scales, the second column is the total standard deviation, the third the t-score and the last column is the two-tailed significance of the difference.

User Experience

The data show that the differences in the user experience for the Alertness as well as the Challenge scales are significant.

Actual movement scored significantly higher on the Alertness scale ($t(19)$=-2.42, p=0.03) which could be attributed to mental tiring process of performing

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
P1	16,90	29,323	,563	,360	,854
P2	16,98	27,717	,809	,739	,816
P3	16,30	26,933	,687	,654	,835
P4	17,65	32,644	,603	,496	,849
P5	16,80	28,574	,696	,623	,833
P6	15,72	32,410	,543	,324	,854
P7	16,65	30,695	,568	,490	,851

Table 3.17: Positive Experiences: Item Total Statistics after corrections

Descriptive Statistics
All Scales

	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
						Std. Error		Std. Error
INegative	1,000	2,800	1,85000	,611727	-,214	,512	-1,280	,992
IPositive	1,143	4,286	2,66429	,811140	-,076	,512	-,207	,992
IFatigue	1,000	4,333	2,38333	1,066639	,202	,512	-1,163	,992
IChallenge	1,200	4,600	2,83000	1,010003	,028	,512	-,770	,992
IImersion	1,333	5,000	3,08333	,904020	,217	,512	,278	,992
IControl	1,000	4,333	2,15000	,951914	,442	,512	-,306	,992
IGoals	1,000	5,000	3,93400	1,000576	-1,266	,512	2,571	,992
IAlertness	1,667	5,000	3,30000	1,174286	,188	,512	-1,405	,992
ANegative	1,000	3,400	1,85000	,745160	,311	,512	-,899	,992
APositive	1,286	4,571	2,90714	,984602	,158	,512	-,833	,992
AFatigue	1,00	3,67	1,9833	,83403	,735	,512	-,685	,992
AChallenge	1,200	4,600	2,43000	,911390	,930	,512	,476	,992
AImmersion	1,667	4,667	3,23333	1,003503	-,104	,512	-1,187	,992
AControl	1,000	4,667	2,45000	1,033220	,390	,512	-,149	,992
AGoals	2,000	5,000	4,11667	,839625	-,880	,512	,485	,992
AAlertness	2,667	5,000	3,95000	,893610	-,223	,512	-1,690	,992

Table 3.18: All scales: Descriptive statistics.

imagined movement. The same trend is also shown in the Fatigue scale, while there is not a significant difference between actual and imagined movement (p=0.12). One possible explanation for this can be found in the correlation between the Fatigue and Alertness scale. These show a strong negative correlation in actual movement (r=-0.707, p<0.001).

Challenge also significantly differs between both kinds of movement ($t(19)$=2.17, p=0.04). User experience data therefore indicates that performing imagined movement is more of a challenge than actual movement is.

One notable remark can be made about Table 3.18: the Goals scale for the imagined movement situation (IGoals) exhibits significant levels of Skewness and Kurtosis. This is the case when the Skewness or Kurtosis statistic is at least twice as high as the standard error of said statistic. Because Student's t-test assumes Gaussianity and this requirement is clearly not met for the IGoals scale this might have an effect when analysing the difference between actual and imagined movement with respect to the Goals scale. However, a

Construction of Scales			
	No. of items	Alpha	Var. explained
Alertness	3	0.783	70.4%
Challenge	5	0.777	56.4%
Control	3	0.783	69.9%
Goals	3	0.754	67.7%
Fatigue	3	0.759	67.6%
Immersion	3	0.620	57.0%
Negative Experiences	5	0.638	41.9%
Positive Experiences	7	0.865	55.8%

Table 3.19: Constructed Scales including α and variance explained by first principal component.

Differences of Imagined vs. Actual Movement				
	Diff of avg	StDev	t	Sig (2-tail)
Alertness	-.65	1.20	-2.42	.03
Challenge	.40	.83	2.17	.04
Control	-.30	1.34	-1.00	.33
Goals	-.18	.50	-1.63	.12
Fatigue	.40	1.11	1.62	.12
Immersion	-.15	.60	-1.12	.28
Negative Experiences	.00	.59	.00	1.00
Positive Experiences	-.24	.89	-1.22	.24

Table 3.20: Paired t-Tests Scales: Difference of the average, Standard Deviation t-score and significance.

non-parametric test for related samples (Wilcoxon Signed Ranks Test) still proved no significant difference in the Goals scales ($W_+(20) = 40$, $p = 0.152$).

Performance

Using the error rate calculated by the classifier from the EEG data we can compare the performance achieved on different subjects. For each subject two error rates are available, one for Actual Movement and one for Imagined Movement. See also Table 3.21 and Table 3.3. The average rate for actual movement is 38.67%, while the average error rate for imagined movement is 42.28%. A Wilcoxon signed-rank test showed that actual movement error rates are significantly lower ($W_+(20) = 48$, $p = 0.0328$). The boxplot in Figure 3.3 visualizes the difference.

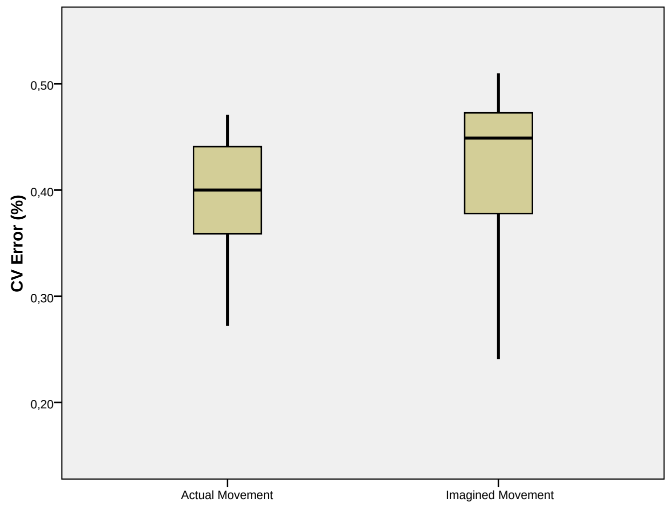

Figure 3.3: Boxplot of the error rate for actual and imagined movement.)

Looking at performance across different demographical groups there are no significant differences between men and women in actual ($t(19)$=0.584, p=0.570) or imagined ($t(19)$=0.205, p=0.840) movement.

Comparing left handed with right handed test subjects also did not show any significant differences in actual or imagined movement ($t(19)$=-0.876, p=0.403 and $t(19)$=0.99, p=0.923 respectively).

Looking at the actual movement performance plotted versus age (see Figure 3.4) and the respective Pearson product-moment correlation coefficient (equivalent to a least squares fit) one can see that there is an apparent correlation between the two (r=0.462, p=0.041). A regression test shows that age explains the variance in Cross Validation Error for 21.3% (R^2=0.213, F=4.821

Descriptive Statistics

	N	Minimum	Maximum	Sum	Mean	Std. Deviation
Actual Movement Performance	20	,1708	,4708	7,7333	,386665	,0722835
IImagined Movement Performance	20	,2408	,5099	8,4554	,422770	,0762328

Table 3.21: Descriptive statistics of error rate.

Cross Validation Errors		
Subject	Imagined Movement	Actual Movement
1	0.4732	0.3847
2	0.4208	0.4380
3	0.4901	0.4444
4	0.4859	0.3803
5	0.4972	0.4592
6	0.2597	0.4437
7	0.4639	0.3972
8	0.4431	0.4028
9	0.4507	0.3444
10	0.5099	0.4167
11	0.3639	0.4099
12	0.2408	0.1708
13	0.3571	0.3132
14	0.4639	0.4361
15	0.3310	0.4493
16	0.4722	0.4708
17	0.3917	0.3375
18	0.4292	0.3732
19	0.4472	0.2722
20	0.4639	0.3889
average	0.4228	0.3867

Table 3.22: Performance for every subject (Actual and Imagined Movement)

p=0.041) and has therefore a significant influence. This result, though, relies heavily on the three outliers at the far right. Furthermore: *cum hoc ergo propter hoc*; like all statistical tests based on correlations, this does not imply a causation. A similar trend was not found for imagined movement, see Figure 3.5. (R^2=0.000, F=0.002, p=0.963)

To asses if there was a significant learning effect of firstly performing one kind of movement and subsequently performing the other kind of movement, a independent samples t-test was carried out. No significant differences between group A and B were found in actual (p=0.586) or imagined (p=0.317) movement.

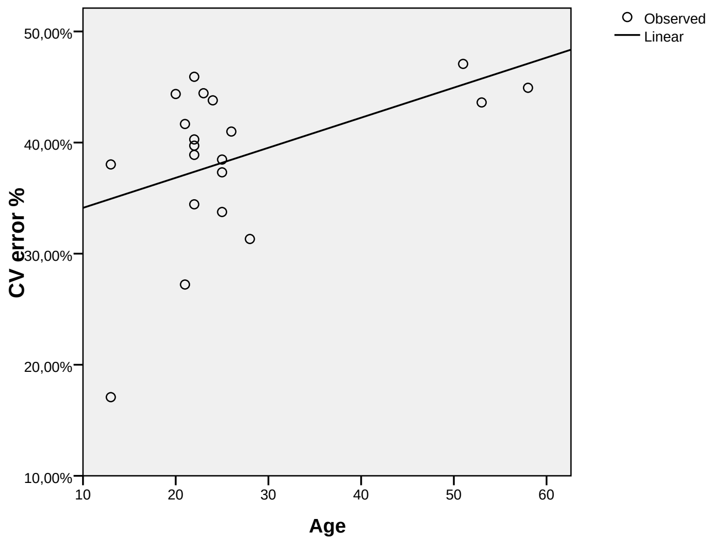

Figure 3.4: Scatterplot of Performance versus Age (Actual Movement)

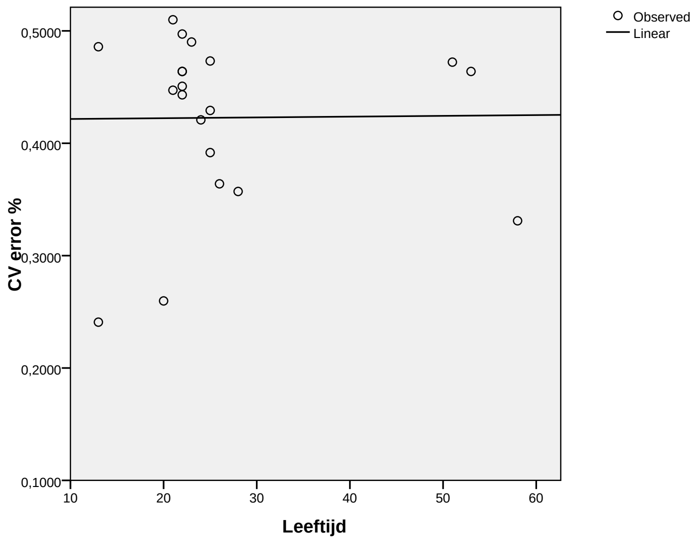

Figure 3.5: Scatterplot of Performance versus Age (Imagined Movement)

4 Discussion

4.1 Discussion of results

Subjects

The experiments were carried out with twenty subjects of which ten were male and ten were female. They were randomly divided in two groups to accommodate the two types of movement in a sequential fashion, averaging out the sequence effects. While there was some difference between the two groups, even if one big group was used, this difference would not have had a significant influence on the results.

An attempt was made to have as much variance in age between the different subjects. While there were two children and three adults, the majority of the subjects were students between 19 and 26 years old. The analysis of age vs. performance relied heavily on the outliers (e.g. the adults). For a definite answer the data from more subjects and especially a more evenly spread out group of subjects should be analysed, for example through stratified sampling [BF].

Because the subject group consisted of equal parts of male and female subjects it was a representative and realistic group. Analyses could be made to determine if there are differences between men and women. There was no trend in accordance with the anecdotal findings that women would perform better than men. The hypothesis that there is no real difference between the performance of men and women is supported by the data.

Questionnaire

The questionnaire that was used based on the GEQ proved a valuable instrument for this study. The Alertness and Control scale were added to the original GEQ questionnaire. The scales in the questionnaire were corrected with the use of the data gathered in this study to create an instrument with a high internal consistency. Before correction the Standardized Alpha of the scales ranged from 0.502 to 0.865 (2 scales α >0.7), after correction the range was from 0.620 to 0.865 (6 scales α >0.7).

While the Control scale did not provide an insight on the differences between actual and imagined movement (possibly there are not any) after correction it proved to be a scale with a high internal consistency (α=0.783) which will be useful for evaluating other BCI games.

The Alertness scale was an important addition to the questionnaire because it measured a significant difference between the two kinds of movement evaluated in this study. With an α of 0.783 it is a scale with a high internal consistency.

User Experience

The corrected questionnaire was used to evaluate the differences in user experience between actual and imagined movement. Hypothesized was that a better performance would lead to a better user experience (especially more positive affect). This however was not proven with significant results. Other significant differences and trends in user experience were found that are discussed below.

Challenge There was a significant (p=0.04) difference in the Challenge scale. Data indicated that imagined movement scores higher on the Challenge scale and therefore requires more effort. It is a task that is harder to accomplish and puts the user under more pressure. Users also indicated this in the remarks they made; some of them did not know how to imagine movement, others found imagined movement more enjoyable than actual movement but also realised that their recognition rate and therefore their game score was lower than with actual movement. This does not have to be a problem if one considers imagined movement as a skill that can be learned and enhanced. While not yet proven, imagined movement might be a skill that can be enhanced by training of the individual.

Alertness Data from the experiment also indicate that there is a significant (p=0.03) difference in Alertness, this time in favour of actual movement. During actual movement users were more focussed, had to do less effort to stay concentrated and were less occupied with other things outside the scope of the game. This is also congruent with the experiences of participants during the experiments, some indicated they almost fell asleep during imagining movement and some felt energised and revived when tapping enthusiastically on the table. This was however not true for everybody, some liked more of a challenge and found actual movement somewhat dull.

Goals While data from the Goals scale did not provide any significant differences, there was a trend to be found (p=0.12). The slight difference in favour of actual movement points in the direction that users knew what the goals of the game were, what to do and how to do it, faster with actual movement than with imagined movement. The fact that some people reported they did not exactly know *how* to imagine movement is probably the underlying factor.

Fatigue The data from the Fatigue scale pointed towards a trend (p=0.12) that imagined movement would be more fatiguing, tiring and requires more concentration. This off course can be related to the result from the Alertness scale. There are correlation between the scales Alertness and Fatigue in actual as well as imagined movement. See also Appendix E: AFatigue and AAlertness (ρ=-0.707, p<0.001) and IFatigue and IAlertness ρ=-0.648, p=0.002). In this case, the tiredness of the user is not a physical one but a form of menta fatigue. The actual movement probably is not enough to physically exert a person enough to overcome this effect.

Performance

The average error rate of all subjects for imagined movement was 42.3% and for actual movement 38.7%. A Wilcoxon test showed that the difference is significant (p=0.0328). This is a support for the hypothesis that actual movement is indeed better recognisable than imagined movement. While there is a significant difference in the average over all subjects, and indeed fifteen participant achieve a lower error rate using actual movement, some individual cases are especially interesting: two participants were able to achieve more than ten percent point increase in performance when using actual movement. This is a very big difference and while someone might seem a BCI illiterate for imagined movement their recognition performance for actual movement is better than average. This could shed a new light on BCI illiteracy for motor imagery. A possible solution might be to let users train with actual movement and gradually shifting to imagined movement, while providing feedback on how well their recognition rate is.

4.2 Significance of Actual Movement

Actual Movement as training

Because of the similarities in brain activity between actual and imagined movement and the somewhat lacking of intuitivity for imagined movement one might suggest using actual movement as a training for using imagined movement. The user of the BCI can get accustomed to using movements for interaction and at the same time trying to imagine the movement. With the acquired data from the actual movement, the imagined movement could be classified. While this hypothesis is not grounded with statistical data, some small scale experiments were carried out during this study. These indicated that it is possible to train a classifier with actual movement and use imagined movement during the game. This methods might therefore be a feasible way to take away part of the awkwardness of just performing imagined movement.

Actual Movement and Motion Tracking

Actual movement can also be used with other interfaces than a BCI. For example an interface such as a motion tracking system, which is probably more reliable at this moment. One big potential advantage of a BCI however is that the measured EEG signals at the brains are always preceding actual muscle movement at the limbs. This advantage is amplified by the onset of a potential in preparation of a movement, the so called Bereitschaftspotential (Readiness Potential) [KD65]. Krauledat et al. showed that this lateralized readiness potential can be used to classify actual movement even before the movement itself is carried out [KDB+04]. This could give a gamer an advantage over other interfaces especially in fast paced, highly reactive games. To this point in time however, the robustness of these methods is lacking as well as algorithms and systems that are fast enough to keep up with these short and hard to detect signals.

4.3 Future Work

Different Ways of Imagining Movement

Future work could include research into the different ways of imagining movement. As McFarland et al. [MMVW00] already explain: when given the instruction to imagine a movement, most people will try to sense the movement. Other kinds of imagination (e.g. visualizing the movement) might activate different cortical areas. Some users might even prefer to visualize a movement if they find it more natural or less tiring. Evaluating the performance and user experience of these different tasks would be a valuable addition to this research.

Improve Questionnaire

The developed questionnaire seems to be an instrument that can aid us in evaluating differences in user experience between different BCI paradigms used with BrainBasher. It might also be of interest for evaluation of BCI games other than BrainBasher. Then further research on the validity and generalizability of the questionnaire is needed.

Subjects

This study used twenty subjects during the experiments for data acquisition, which is *a lot* compared to other BCI studies. Although this might seem a rigorous approach, every study can use more subjects, including this study. This is especially the case because the experiment had a crossover design in which two conditions are compared between two groups. Furthermore, we do not yet know which variables have an influence on BCI performance. Still more research is needed to explain inter and intra-subject variability and inter and intra-session variability.

Materials

Delay

Although the game works in an online manner and the classification algorithm is fast enough to be computed realtime, there always is an inherent delay in feedback. This is due to the fact that the classification algorithm needs a measurement of EEG data of a few seconds. Currently this measurement is two seconds. The consequence is that users get feedback of what they did with a two second delay. This delay sometimes leads to confusion and a lower positive affect towards the system. Future research might include shortening the response time of the underlying system of the game and finding out what this does for acceptance of and positive affection towards the BCI.

Dry Electrodes

One issue that remains when one wants to apply BCI's for casual gaming is the time that is needed to put on an EEG cap and connect the equipment. A skilled experimenter can set up an 32 electrode EEG cap within ten minutes, but this still is unacceptably long for a session of casual gaming. Especially when a consecutive training session is needed every time the EEG cap is set up, most players will lose interest. Future EEG headsets that incorporate (semi) dry electrodes and a wireless connection to a computer will boost the acceptance of BCI gaming. Furthermore research into zero training or a form of hybrid training whilst gaming at the same time should be conducted.

Good vs. bad recognition

One aspect of the game that should be different in the future is handling the subjects that cannot get the BCI to recognize their brain activity. Two participants could not get the BCI to recognize any action they made, the classifier just did not produce a classification confidence over 60% for either left hand or right hand movement. An example of such a game session can be seen in Figure 4.1. The progress bar shows more or less a flat line without definite peaks for either colour, e.g. the classifier does not make strong decisions. In this typical example the cause is probably a dominant ‘rest’ class. The lower area in the progress bar (grey) is already around 50%, therefore making it nearly impossible for either ‘hand’ class to achieve a 60% level.

A gamescreen from a person that achieves a better recognition can be seen in Figure 4.2. Clearly can be seen that the progress bar shows high peaks of blue and purple, indicating that the classifier makes strong decisions.

This 'flatlining' (pun definitely intended) makes playing the game very boring. Users are probably only completing the game because it is an experiment and do not want to disappoint the experimenter (compliance bias). A

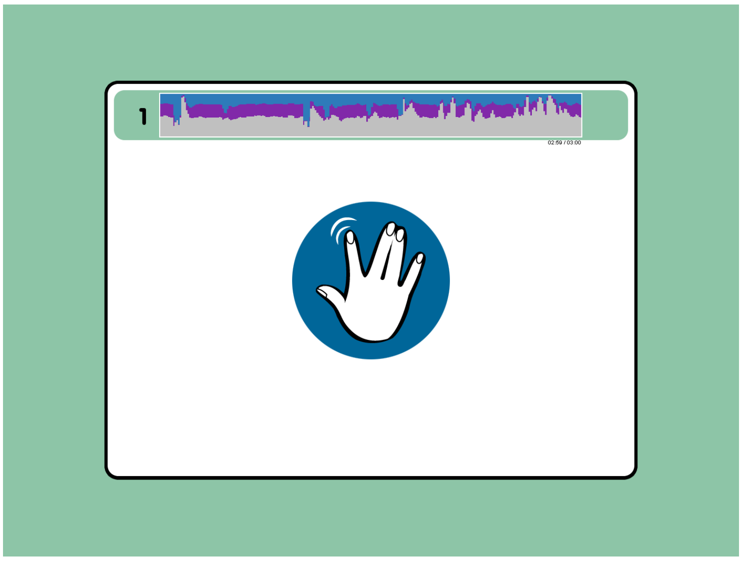

Figure 4.1: Typical screen of a badly recognized subject.

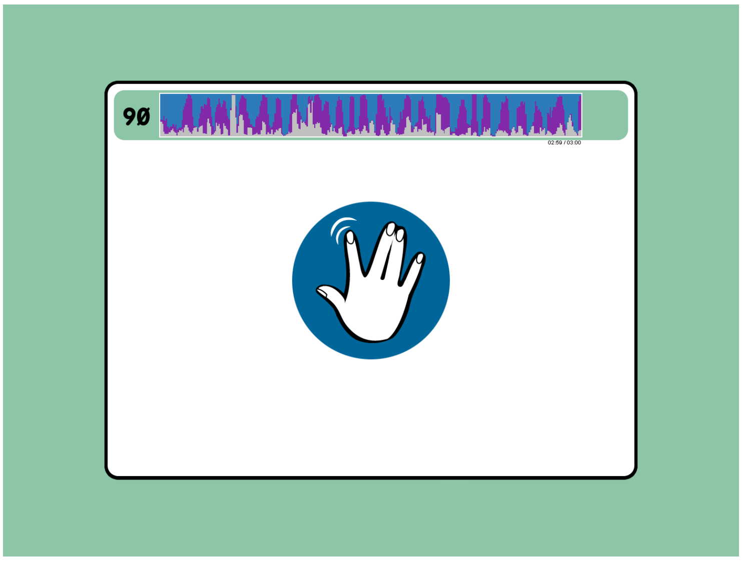

Figure 4.2: Typical screen of a well recognized subject.

solution of this problem could consist of using another classification method when said flatlining is detected, or shortening/altering the game for example.

5

Conclusions

Results from this research showed that differences in user experience and in performance between actual and imagined movement in BCI gaming do exist. Actual movement produces a more reliable signal while the user stays more alert. On the other hand, imagined movement is more challenging.

To be able to assess the differences in user experience between actual and imagined movement, a special version of the GEQ for BCI games was developed. While this questionnaire was found to be a numerically grounded tool to be used in this setting, further research for validation is needed. User experience data from this questionnaire showed two significant differences. Users found more challenge in performing imagined movement. This might be due to a higher error rate, which makes sense; looking at the average error rate, it is harder to perform imagined movement. If we assume imagined movement is a skill that can be learned this might be an advantage for using imagined movement. Gamers are always looking for challenges and limitations that they can overcome by practice [NTP+08].

On the other hand, for a few test subjects, the BCI system could not correctly recognize any movement. This corresponds to an error rate of 50%, in which case simple random 'guessing' would be as good as any classification. Participants who achieved a high error rate also were not able to score any game points (other than maybe by chance). This is an issue that frustrates the user and is something that has to be resolved for wider acceptance of BCI gaming. The problem of not being able to be *understood* by a BCI which is referred to as BCI illiteracy [SBM08].

Alertness is the other scale in which a difference was found. This alertness has to do with the state of mind of the user after they played the game. The fact that they felt less alert after performing imagined movement is explainable. Imaginary movement requires more concentration and is a less natural action to perform. Doing something you do everyday does not tire you as much as doing something completely new. This was also reflected in the Fatigue scale, which scored slightly higher for imagined movement.

The generalizability over various demographic groups was good and there were no significant differences in performance. While there have been some anecdotal findings that women would be better in communicating through a BCI, results show no significant differences between men and women. Data also did not show any differences between left and right handed people.

While the gathered data does not provide a clear view on how age is related to performance in the game, one might hypothesize that imagined movement is a skill of young children who mimic movements of others. A child sees someone performing a certain movement that can be of advantage to the child, for example grabbing something, they then try to perform it themselves. This probably is a skill that fades over time when a person gets older. While at a higher age humans are still able to mimic movements, it takes them more time to learn it. This is possibly a ground for older people not performing to well at imagined movement. This was also reported by test subjects to the experimenter. They do not know how or what they should imagine.

Bibliography

[BF] P. J. Bickel and D. A. Freedman. Asymptotic Normality and the Bootstrap in Stratified Sampling. volume 12.

[BPK+] Benjamin Blankertz, Florin Popescu, Matthias Krauledat, Siamac Fazli, Michael Tangermann, and KlausRobert Mller.

[BR08] D. Oude Bos and B. Reuderink. Brainbasher: a bci game. In P. Markopoulos, J. Hoonhout, I. Soute, and J. Read, editors, *Extended Abstracts of the International Conference on Fun and Games 2008, Eindhoven, Netherlands*, pages 36–39, Eindhoven, October 2008. Eindhoven University of Technology.

[Cor93] J.M. Cortina. What is Coefficient Alpha? An Examination of Theory and Applications. *Journal of Applied Psychology*, 78(1):98–104, 1993.

[Cro51] L.J. Cronbach. Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3):297–334, 1951.

[CV77] R. B. Catell and S. Vogelmann. A Comprehensive Trial Of The Scree And Kg Criteria For Determining The Number Of Factors. *Multivariate Behavioral Research*, 12(3):289–325, 1977.

[DW08] H. Desurvire and C. Wiberg. Evaluating User Experience and Other Lies in Evaluating Games. page 1, 2008.

[GAG07] B. Graimann, B. Allison, and A. Gräser. New Applications for Non-invasive Brain-Computer Interfaces and the Need for Engaging Training Environments. *BRAINPLAY 07 BCI and Games Workshop at ACE*, 2007.

[Hea08] W. Head. Nintendo wiining the console war, 2008. `http://www.igizmo.co.uk/articles/news/744-gaming-nintendo-wiining-console-war`.

[IdKP+07] W. IJsselsteijn, Y. de Kort, K. Poels, A. Jurgelionis, and F. Bellotti. Characterising and Measuring User Experiences in Digital Games. *Intern. Conf. on Advances in Computer Entertainment Technology*, 2007.

[Jea94] M. Jeannerod. The representing brain Neural correlates of motor intention and imagery. *Behavioural Brain Sciences*, 17:187–245, 1994.

[KBL+07] T. Kayagil, O. Bai, P. Lin, S. Furlani, S. Vorbach, and M. Hallett. Binary EEG Control for Two-Dimensional Cursor Movement: An Online Approach. *IEEE/ICME International Conference on Complex Medical Engineering*, pages 1542–1545, May 2007.

[KD65] H. Kornhuber and L. Deecke. Hirnpotentialänderungen bei willkürbewegungen und passiven bewegungen des menschen: Bereitschaftspotential und reafferente potentiale. *Pflügers Archiv European Journal of Physiology*, 284(1):1–17, 1965.

[KDB+04] M. Krauledat, G. Dornhege, B. Blankertz, F. Losch, G. Curio, and K.-R. Müller. IMPROVING SPEED AND ACCURACY OF BRAIN-COMPUTER INTERFACES USING READINESS POTENTIAL FEATURES. *Proceedings of the 26th Annual International Conference of the IEEE EMBS*, 26:4511–4515, 2004.

[Kol91] Zoltan Joseph Koles. The quantitative extraction and topographic mapping of the abnormal components in the clinical eeg. *Electroencephalography and Clinical Neurophysiology*, 79(6):440–447, December 1991.

[Lar08] J. Larsen. Evaluating User Experience - how game reviewers do it. page 1, 2008.

[LKF+05] Robert Leeb, Claudia Keinrath, Doron Friedman, Christoph Guger, Christa Neuper, Maia Garau, Angus Antley, Anthony Steed, Mel Slater, and Gert Pfurtscheller. Walking from thoughts: Not the muscles are crucial but the brain waves! In *Proceedings of the 8th Annual International Workshop on Presence*, pages 25–32, 2005.

[LLK+07] R. Leeb, F. Lee, C. Keinrath, R. Scherer, H. Bischof, and G. Pfurtscheller. Brain–Computer Communication: Motivation, Aim, and Impact of Exploring a Virtual Apartment. *Neur. Sys. and Rehab. Eng., IEEE Trans. on*, 15(4):473–482, 2007.

[LLR+08] A. Lecuyer, F. Lotte, R.B. Reilly, R. Leeb, M. Hirose, and M. Slater. Brain-computer interfaces, virtual reality, and videogames. *Computer*, 41(10):66–72, Oct. 2008.

[MMVW00] D.J. McFarland, L.A. Miner, T.M. Vaughan, and J.R. Wolpaw. Mu and Beta Rhythm Topographies During Motor Imagery and Actual Movements. *Brain Topography*, 12(3):177–186, 2000.

[NHJC08] Vadim V. Nikulin, Friederike U. Hohlefeld, Arthur M. Jacobs, and Gabriel Curio. Quasi-movements: A novel motorcognitive phenomenon. *Neuropsychologia*, 46:727–742, 2008.

[NTA+08] Anton Nijholt, Desney Tan, Brendan Allison, Jose del R. Milan, and Bernhard Graimann. Brain-computer interfaces for hci and games. In *CHI '08: CHI '08 extended abstracts on Human factors in computing systems*, pages 3925–3928, New York, NY, USA, 2008. ACM.

[NTP+08] A. Nijholt, D. Tan, G. Pfurtscheller, C. Brunner, J.R. Millán, B. Allison, B. Graimann, F. Popescu, B. Blankertz, and K.R. Müller. Brain-Computer Interfacing for Intelligent Systems. *IEEE Intell. Systems*, 23(3):72–79, 2008.

[Pfu01] G. Pfurtscheller. Functional brain imaging based on erd/ers. *Vision Research*, 41(10-11):1257 – 1260, 2001.

[PSVH03] Jaime A. Pineda, David S. Silverman, Andrey Vankov, and John Hestenes. Learning to control brain rhythms: making a brain-computer interface possible. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2):181–184, june 2003.

[SBM08] C. Sanelli, M. Braun, M. Tangermann, and K. Mller. Estimating noise and dimensionality in bci data sets: Towards bci illiteracy comprehension. *Proceedings of the 4th International Brain-Computer Interface Workshop and Training Course 2008*, 2008.

6 Appendices

6.1 Appendix A - Informed Consent

Informed consent

Dear participant,

You have volunteered to participate in this study. Here you will find some information about your rights and the procedure of the following experiment. Please read the following sections carefully.

1) The purpose of the study

The aim of this study is to provide new knowledge about experience of users playing a game in a Brain Computer Interaction environment.

2) The procedure of the experiment

The purpose of this experiment is to complete two tasks. Both tasks consist of achieving a score as high as possible. The tasks differ from each other in the way you will steer the computer. Both tasks consist of several different actions. Before both tasks begin you will get the possibility to practice. (The experimenter will give you further instructions when this is needed.)
At the beginning of the experiment the electrode cap is set up. This process takes about 15-20 minutes. The actual experiment takes about 40 minutes. The experimenter will be with you in the experimental space throughout the experiment and is available for questions at any time.
After the experiment you will learn more about the meaning and purpose of this investigation. Please do not give this information to others, so that the objectivity of any potential test subjects is preserved.

3) Risks and side effects

This study is based on present knowledge of the investigator in charge and is safe and painless for the participants. By participating in this study you will not undergo specific risks, and there are no side effects reported. *Since this study as a whole is new, the occurrence of unknown side effects, however, cannot be excluded.*

Important: Please inform the investigator as soon as possible, if you have illnesses or are undergoing medical treatment. Let the investigator immediately know if you ever had an epileptic seizure *or suffer from tinnitus*. For questions about this please contact the investigator.

4) Termination of the experiment

You have the right to cancel at any time and without stating the ground. The participation is entirely voluntary and without obligations. There are no disadvantages by suspending the investigation.
During the experiment, you have multiple opportunities for a break. If you ever want a break or even need to go to the toilet, this is possible at any time.
Should you at any time during the experiment experience discomfort, please inform the investigator immediately.

5) Confidentiality

The provisions of data protection are respected. Personal data will not be disclosed to third parties. The collected data is anonymised by us and only in this form processed and published.

6) Declaration

Please confirm to the following statement by placing your signature.
"I, the undersigned, hereby state that the investigator of this study informed me on the points above. I have read and understood this statement. I agree with each of the points. I hereby authorize that the data acquired in this investigation is analysed for scientific purposes and anonymously published. I was informed on my rights as a test subject and on the voluntary participation in this study."

…………………………………...
Name

…………………………………………………………………..…………………
Place Date Signature

………………………………………
If the participant is a minor, signature of a parent is required.

Verklaring van toestemming

Beste deelnemer,

U hebt toegezegd vrijwillig mee te doen aan een onderzoek. In dit document staat informatie over uw rechten en de procedure van het experiment. Lees de volgende paragrafen met veel aandacht.

1) Doel van het onderzoek

Het doel van dit onderzoek is meer inzicht te verkrijgen in de ervaring van gebruikers met het spelen van een spel in een Brein Computer Interactie omgeving.

2) Procedure tijdens het experiment

In dit experiment is het de bedoeling dat u twee taken voltooid. Beide taken bestaan uit het behalen van een zo hoog mogelijke score in het spel. De ene taak verschilt van de ander in de manier van aansturing door u. Beide taken bestaan uit een tiental acties. U krijgt voorafgaand aan elke taak de tijd en mogelijkheid om te oefenen. (De onderzoeker zal u verdere uitleg geven over het experiment wanneer dit nodig is.)
Voorafgaand aan het eigenlijke experiment wordt de muts met elektrodes bij u opgezet. Deze handeling duurt ongeveer 15 tot 20 minuten. Het eigenlijke experiment duurt ongeveer veertig minuten. De onderzoeker zal gedurende het gehele experiment bij u zijn en beschikbaar zijn voor vragen en opmerkingen. Na afloop van het experiment krijgt u meer te horen over de achtergrond van het onderzoek en de manier waarop de resultaten gebruikt gaan worden. Speel deze informatie alstublieft niet door aan anderen zodat de objectiviteit en onbevooroordeeldheid van potentiële proefpersonen niet verloren gaat.

3) Risico's en neveneffecten

Het experiment is gebaseerd op aanwezige kennis van de onderzoeker en is veilig en pijnloos voor de proefpersonen. Door mee te doen aan dit experiment zal u geen risico's ondervinden. Daarnaast zijn er geen neveneffecten bij dit soort experimenten gemeld. *Omdat dit onderzoek in zijn geheel nieuw is, zou het echter kunnen zijn dat er onbekende neveneffecten kunnen optreden, de kans hierop is echter zeer klein.*

Belangrijk: Indien u aan een (chronische) ziekte lijdt of onder medische behandeling bent, laat het de onderzoeker dan zo snel mogelijk weten. Ook als u ooit een epileptische aanval het gehad. Als u hierover vragen hebt schroom niet om de onderzoeker hierover aan te spreken.

4) Beëindiging van het experiment

U hebt het recht om op elk moment het experiment te beëindigen zonder opgaaf van reden. De deelname is volledig vrijwillig en zonder verplichtingen. Er zijn geen nadelen voor u verbonden aan beëindiging van het experiment uwerzijds.
Tijdens het experiment hebt u meerdere mogelijkheden om een pauze te nemen. Als u tijdens het experiment een pauze wilt of mogelijk het toilet wilt bezoeken is dat mogelijk op elk moment. Als u (op welk moment dan ook) zich niet op uw gemak voelt, informeer dan onmiddellijk de onderzoeker.

5) Privacy en discretie

De beperkingen en regels die van toepassing zijn op de geheimhouding van de data worden gerespecteerd. Persoonlijke kenmerken zullen nooit ofte nimmer doorgespeeld worden aan derden. De verzamelde data wordt geanonimiseerd en zal enkel in voorgenoemde vorm geanalyseerd en gepubliceerd worden.

6) Verklaring

Door hieronder uw handtekening te zetten verklaart u het eens te zijn met onderstaande verklaring.

"Ik, de ondertekenende, verklaar hierbij dat de onderzoeker me geïnformeerd heeft over bovenstaande punten. Ik heb de verklaring van toestemming gelezen en begrepen. Ik ben het eens met alle bovenstaande punten. Ik geef bij dezen toestemming om de te verzamelen data in geanonimiseerde vorm te laten analyseren en publiceren voor wetenschappelijke doeleinden. Ik ben op de hoogte gesteld van mijn rechten als proefpersoon en mijn vrijwillige participatie in dit onderzoek."

……………….……………………….
Naam

…………………………………………………………………………………...
Plaats Datum Handtekening

Als de proefpersoon een minderjarige is, is ook de handtekening van een ouder of wettelijke voogd vereist.

……………….……………………….
Handtekening

6.2 Appendix B - User information survey

Informatie Proefpersoon

Algemene informatie experiment *in te vullen door onderzoeker*
Datum: ___ - ___ - _______ *Start:* ___ : ___ *Eind:* ___ : ___
Omschrijving: ____

____

____

Software: ____
Actiview settings: *Bandpass:* ____ Hz - ____ Hz *Sample frequency:* _____ Hz
Experiment definitie: ____
Onderzoeker(s): ____

Algemene informatie proefpersoon

in te vullen door proefpersoon	*in te vullen door onderzoeker*
Geslacht: m / v	*ID:*
Leeftijd: ____	*Hoofd omtrek*:___ cm *Cap*: _____
Dominantie: linkshandig / rechtshandig	*Nasion-inion:* ___ cm *Cz*: ___cm
Opleiding: ___________________________	*Slaap-slaap:* ___ cm *Cz*: ___cm
Beroep: _____________________________	*Haaromschrijving:* _____________
	Haarproducten: ________________

Visuele hulpmiddelen: o contactlenzen o bril o geen

Alcoholconsumptie: o meerdere dagelijks o dagelijks o wekelijks
o maandelijks o nooit
Voor experiment: ___________________________

Koffieconsumptie: o 5+ kopjes per dag o 3-5 kopjes o 1-3 kopjes
o minder dan 1 kopje o nooit
Voor experiment: ___________________________

Zwarte/groene thee consumptie: o 5+ kopjes per dag o 3-5 kopjes o 1-3 kopjes
o minder dan 1 kopje o nooit
Voor experiment: ___________________________

Frisdrank (cafeïnehoudend) Consumptie: o 5+ glazen per dag o 3-5 glazen o 1-3 glazen
o minder dan 1 glas o nooit
Voor experiment: ___________________________

Tabakconsumptie: o 2+ pakjes per dag o 1-2 per dag o 1 pakje per dag
o minder dan 1 pakje o af en toe o nooit
Voor experiment: ___________________________

Aantal uren slaap: per nacht: ___ voor experiment: ___

Mate van alertheid op dit moment: niet alert o o o o o zeer alert

Medicijnen: ____

Aandachts-/neurologische/psychiatrische problemen:

____

Werken met de PC	o 6+ uur per dag	o 4-6 uur p.d.	o 2-4 uur p.d.
(niet per se voor beroep)	o dagelijks, <2 uur	o wekelijks	o maandelijks
	o minder dan maandelijks		o nooit
Computerspellen spelen	o 6+ uur per dag	o 4-6 uur p.d.	o 2-4 uur p.d.
	o dagelijks, < 2 uur	o wekelijks	o maandelijks
	o minder dan maandelijks		o nooit

Handigheidstest:

Geef aan met welke hand je de volgende taken zou uitvoeren:

Schrijven	Links / Rechts
Tekenen	Links / Rechts
Bal gooien	Links / Rechts
Knippen	Links / Rechts
Tanden poetsen	Links / Rechts
Brood snijden met mes	Links / Rechts
Eten met een lepel	Links / Rechts
Een computermuis gebruiken	Links / Rechts
Een lucifer aansteken *(de andere hand houdt het doosje vast)*	Links / Rechts
Een pot openmaken *(de andere hand houdt de pot vast)*	Links / Rechts

Met welke hand schrijven:

Moeder	L / R / ?	*Zus*	L / R / ?	*Dochter*	L / R / ?
Vader	L / R / ?	*Broer*	L / R / ?	*Zoon*	L / R / ?

Eventuele extra broers/zussen/dochters/zonen:

L / R	L / R	L / R	L / R
L / R	L / R	L / R	L / R

Opmerkingen:

____
____
____
____
____
____
____

6.3 Appendix C - Administered Questionnaire

Gebruikerservaringen Imagined Movement

Geef voor elk van onderstaande items aan hoe je je voelde tijdens of na het spelen. Gebruik hierbij de volgende schaal:

Schaalverdeling:

Helemaal oneens	Enigszins oneens	Neutraal	Enigszins eens	Helemaal eens
O	O	O	O	O

Ik vond het leuk om het spel te spelen.	oneens	O O O O O	eens
Ik voelde me energieker.	oneens	O O O O O	eens
Ik vond het tijdverspilling.	oneens	O O O O O	eens
Ik ben uitgeput.	oneens	O O O O O	eens
Ik vond het saai.	oneens	O O O O O	eens
Ik had snel door hoe ik het spel moest spelen.	oneens	O O O O O	eens
Ik voelde me trots.	oneens	O O O O O	eens
Ik kreeg hoofdpijn tijdens het spelen.	oneens	O O O O O	eens
Ik heb spijt van het spelen.	oneens	O O O O O	eens
Ik voelde me onder druk gezet.	oneens	O O O O O	eens
Ik vond het moeilijk om terug te keren naar de werkelijkheid.	oneens	O O O O O	eens
Ik voel me er slecht over.	oneens	O O O O O	eens
Ik voelde me gedesoriënteerd.	oneens	O O O O O	eens
Ik kon me goed concentreren.	oneens	O O O O O	eens
Ik raakte geïrriteerd.	oneens	O O O O O	eens
Ik had het gevoel dat de computer mijn bewegingen herkende.	oneens	O O O O O	eens
Ik had het gevoel dat ik betere dingen had kunnen doen.	oneens	O O O O O	eens
Ik voelde me voldaan.	oneens	O O O O O	eens
Ik vond het indrukwekkend.	oneens	O O O O O	eens
Het voltooien van de taak voelde als een overwinning.	oneens	O O O O O	eens
Ik was gefrustreerd.	oneens	O O O O O	eens
Ik begreep wat ik moest doen.	oneens	O O O O O	eens
Ik was met andere zaken bezig.	oneens	O O O O O	eens
Het design van het spel sprak me niet aan.	oneens	O O O O O	eens
Ik had het gevoel dat de computer me feilloos aanvoelde.	oneens	O O O O O	eens
Ik had het gevoel dat het spel mij bestuurde.	oneens	O O O O O	eens
Ik moest er veel moeite in steken.	oneens	O O O O O	eens
Het spel heeft een steile leercurve.	oneens	O O O O O	eens
Ik besefte hoe goed ik presteerde tijdens het spel.	oneens	O O O O O	eens
Goed scoren werd gedurende het spel steeds makkelijker.	oneens	O O O O O	eens
Ik voelde me machtig.	oneens	O O O O O	eens
Ik voelde me uitgedaagd.	oneens	O O O O O	eens
Ik moest moeite doen om me te concentreren.	oneens	O O O O O	eens
Ik vond het moeilijk.	oneens	O O O O O	eens
Ik voel me vermoeid.	oneens	O O O O O	eens
Ik voelde me tevreden.	oneens	O O O O O	eens
Het doel van het spel is duidelijk.	oneens	O O O O O	eens
Ik had het gevoel dat ik terugkwam van een reis.	oneens	O O O O O	eens
Ik had het gevoel dat ik het spel bestuurde.	oneens	O O O O O	eens
Ik kon me goed focussen op de acties die ik moest doen.	oneens	O O O O O	eens
Ik was mijn gevoel voor tijd kwijt.	oneens	O O O O O	eens
Ik vond het een nieuwe ervaring.	oneens	O O O O O	eens

Gebruikerservaringen Actual Movement

Geef voor elk van onderstaande items aan hoe je je voelde tijdens of na het spelen. Gebruik hierbij de volgende schaal:

Schaalverdeling:

Helemaal oneens	Enigszins oneens	Neutraal	Enigszins eens	Helemaal eens
O	O	O	O	O

Ik vond het leuk om het spel te spelen.	oneens	O O O O O	eens
Ik voelde me energieker.	oneens	O O O O O	eens
Ik vond het tijdverspilling.	oneens	O O O O O	eens
Ik ben uitgeput.	oneens	O O O O O	eens
Ik vond het saai.	oneens	O O O O O	eens
Ik had snel door hoe ik het spel moest spelen.	oneens	O O O O O	eens
Ik voelde me trots.	oneens	O O O O O	eens
Ik kreeg hoofdpijn tijdens het spelen.	oneens	O O O O O	eens
Ik heb spijt van het spelen.	oneens	O O O O O	eens
Ik voelde me onder druk gezet.	oneens	O O O O O	eens
Ik vond het moeilijk om terug te keren naar de werkelijkheid.	oneens	O O O O O	eens
Ik voel me er slecht over.	oneens	O O O O O	eens
Ik voelde me gedesoriënteerd.	oneens	O O O O O	eens
Ik kon me goed concentreren.	oneens	O O O O O	eens
Ik raakte geïrriteerd.	oneens	O O O O O	eens
Ik had het gevoel dat de computer mijn bewegingen herkende.	oneens	O O O O O	eens
Ik had het gevoel dat ik betere dingen had kunnen doen.	oneens	O O O O O	eens
Ik voelde me voldaan.	oneens	O O O O O	eens
Ik vond het indrukwekkend.	oneens	O O O O O	eens
Het voltooien van de taak voelde als een overwinning.	oneens	O O O O O	eens
Ik was gefrustreerd.	oneens	O O O O O	eens
Ik begreep wat ik moest doen.	oneens	O O O O O	eens
Ik was met andere zaken bezig.	oneens	O O O O O	eens
Het design van het spel sprak me niet aan.	oneens	O O O O O	eens
Ik had het gevoel dat de computer me feilloos aanvoelde.	oneens	O O O O O	eens
Ik had het gevoel dat het spel mij bestuurde.	oneens	O O O O O	eens
Ik moest er veel moeite in steken.	oneens	O O O O O	eens
Het spel heeft een steile leercurve.	oneens	O O O O O	eens
Ik besefte hoe goed ik presteerde tijdens het spel.	oneens	O O O O O	eens
Goed scoren werd gedurende het spel steeds makkelijker.	oneens	O O O O O	eens
Ik voelde me machtig.	oneens	O O O O O	eens
Ik voelde me uitgedaagd.	oneens	O O O O O	eens
Ik moest moeite doen om me te concentreren.	oneens	O O O O O	eens
Ik vond het moeilijk.	oneens	O O O O O	eens
Ik voel me vermoeid.	oneens	O O O O O	eens
Ik voelde me tevreden.	oneens	O O O O O	eens
Het doel van het spel is duidelijk.	oneens	O O O O O	eens
Ik had het gevoel dat ik terugkwam van een reis.	oneens	O O O O O	eens
Ik had het gevoel dat ik het spel bestuurde.	oneens	O O O O O	eens
Ik kon me goed focussen op de acties die ik moest doen.	oneens	O O O O O	eens
Ik was mijn gevoel voor tijd kwijt.	oneens	O O O O O	eens
Ik vond het een nieuwe ervaring.	oneens	O O O O O	eens

Wat vond je van de training voorafgaand aan het spel?

Wat vond je leuker om te doen, imaginaire of echte beweging?

Wat vond je van het spel?

Heb je nog andere opmerkingen of suggesties? (Kan je ook direct aan de onderzoeker vertellen.)

6.4 Appendix D - Protocol for Experiments

User side steps

1. The experimenter starts BrainBasher on the game computer.
2. Subject comes into the room and is welcomed by the experimenter.
3. The subject is offered to sit behind the game computer and to adjust his chair to his own comfort.
4. The experimenter explains the procedure regarding the experiment and asks the subject to sign the Informed Consent form.
5. The experimenter asks the subject to fill out a form asking general information about the subject. While the subject is filling out the form, the experimenter starts with setting up the EEG cap.
6. The measures of the skull of the subject are taken and the appropriate size of EEG cap is put on and lined out.
7. While injecting the electro-gel into the holes of the EEG cap the experimenter continuously explains what he is doing.
8. After the EEG cap is set up, check de off sets in ActiView, off sets should be <= 25. Correct off sets by inserting extra electro gel.
9. Explain the game to the user with help of the help screens. Instruct explicitly the movement: "While laying your wrist on the table, tap with your whole hand, with stretched fingers."
10. After the game, take off the EEG cap and offer the subject to shower including shampoo and towel.

System side steps

1. Start ActiView (The EEG recording software).
2. Load the appropriate configuration file.
3. Start a new BDF file in ActiView.
4. Connect with PuTTY to an HMI server.
5. Turn on logging within PuTTY.
6. Go to the appropriate directory on the server and start MatLab.
7. Start the BrainBasher backend.

8. The system is now ready, if the user is also ready, the first training session can begin.

9. After the first training session, the classifier will computer intermediate results, do nothing.

10. After two training sessions the system will again compute performance results. After this it is ready for a game session.

11. When the game session has ended, restart the classifier by interrupting the MatLab routine and restarting it.

12. During the second part of the experiment follow the same protocol as during the first.

6.5 Appendix E - Scale Correlations

Correlations

		INegative	IPositive	IFatigue	IChallenge	IImersion	IControl	IGoals	IAlertness
INegative	rho	1,000	-,470*	,668**	,819**	-,154	-,628**	-,475*	-,745**
	p		,036	,001	,000	,517	,003	,034	,000
IPositive	rho	-,470*	1,000	-,342	-,141	,252	,448*	,379	,372
	p	,036		,140	,553	,284	,047	,100	,107
IFatigue	rho	,668**	-,342	1,000	,399	,238	-,405	-,418	-,648**
	p	,001	,140		,081	,312	,076	,067	,002
IChallenge	rho	,819**	-,141	,399	1,000	-,172	-,443	-,334	-,641**
	p	,000	,553	,081		,468	,051	,150	,002
IImersion	rho	-,154	,252	,238	-,172	1,000	,270	,208	,278
	p	,517	,284	,312	,468		,249	,380	,235
IControl	rho	-,628**	,448*	-,405	-,443	,270	1,000	,355	,570**
	p	,003	,047	,076	,051	,249		,124	,009
IGoals	rho	-,475*	,379	-,418	-,334	,208	,355	1,000	,196
	p	,034	,100	,067	,150	,380	,124		,406
IAlertness	rho	-,745**	,372	-,648**	-,641**	,278	,570**	,196	1,000
	p	,000	,107	,002	,002	,235	,009	,406	
ANegative	rho	,636**	-,132	,244	,613**	-,449*	-,343	-,400	-,499*
	p	,003	,578	,300	,004	,047	,139	,081	,025
APositive	rho	-,229	,522*	,021	-,082	,553*	,082	,298	,238
	p	,331	,018	,929	,732	,012	,730	,201	,313
AFatigue	rho	,242	,032	,343	,251	-,029	-,137	-,379	-,186
	p	,303	,893	,139	,287	,903	,566	,099	,433
AChallenge	rho	,405	-,055	-,067	,635**	-,514*	-,212	-,178	-,432
	p	,077	,818	,780	,003	,020	,370	,452	,057
AImmersion	rho	-,014	,018	,322	-,128	,809**	-,008	-,001	,141
	p	,952	,939	,166	,589	,000	,973	,996	,553
AControl	rho	-,198	,339	-,085	-,209	,646**	,094	,438	,312
	p	,402	,143	,721	,377	,002	,693	,054	,181
AGoals	rho	-,319	,432	-,209	-,195	,349	,284	,866**	,099
	p	,170	,057	,376	,411	,132	,225	,000	,678
AAlertness	rho	-,361	,287	-,476*	-,134	,201	,277	,525*	,349
	p	,118	,220	,034	,572	,396	,236	,017	,131

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Correlations

		ANegative	APositive	AFatigue	AChallenge	AImmersion	AControl	AGoals	AAlertness
INegative	rho	,636**	-,229	,242	,405	-,014	-,198	-,319	-,361
	p	,003	,331	,303	,077	,952	,402	,170	,118
IPositive	rho	-,132	,522*	,032	-,055	,018	,339	,432	,287
	p	,578	,018	,893	,818	,939	,143	,057	,220
IFatigue	rho	,244	,021	,343	-,067	,322	-,085	-,209	-,476*
	p	,300	,929	,139	,780	,166	,721	,376	,034
IChallenge	rho	,613**	-,082	,251	,635**	-,128	-,209	-,195	-,134
	p	,004	,732	,287	,003	,589	,377	,411	,572
IImersion	rho	-,449*	,553*	-,029	-,514*	,809**	,646**	,349	,201
	p	,047	,012	,903	,020	,000	,002	,132	,396
IControl	rho	-,343	,082	-,137	-,212	-,008	,094	,284	,277
	p	,139	,730	,566	,370	,973	,693	,225	,236
IGoals	rho	-,400	,298	-,379	-,178	-,001	,438	,866**	,525*
	p	,081	,201	,099	,452	,996	,054	,000	,017
IAlertness	rho	-,499*	,238	-,186	-,432	,141	,312	,099	,349
	p	,025	,313	,433	,057	,553	,181	,678	,131
ANegative	rho	1,000	-,565**	,538*	,766**	-,486*	-,569**	-,419	-,497*
	p		,009	,014	,000	,030	,009	,066	,026
APositive	rho	-,565**	1,000	-,176	-,488*	,629**	,817**	,435	,353
	p	,009		,458	,029	,003	,000	,055	,126
AFatigue	rho	,538*	-,176	1,000	,236	-,163	-,262	-,390	-,707**
	p	,014	,458		,316	,493	,264	,090	,000
AChallenge	rho	,766**	-,488*	,236	1,000	-,537*	-,574**	-,257	-,179
	p	,000	,029	,316		,015	,008	,274	,450
AImmersion	rho	-,486*	,629**	-,163	-,537*	1,000	,610**	,223	,183
	p	,030	,003	,493	,015		,004	,345	,439
AControl	rho	-,569**	,817**	-,262	-,574**	,610**	1,000	,462*	,418
	p	,009	,000	,264	,008	,004		,040	,066
AGoals	rho	-,419	,435	-,390	-,257	,223	,462*	1,000	,499*
	p	,066	,055	,090	,274	,345	,040		,025
AAlertness	rho	-,497*	,353	-,707**	-,179	,183	,418	,499*	1,000
	p	,026	,126	,000	,450	,439	,066	,025	

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).